*Journal of Phonetics* (2002) **30,** 229–258
doi:10.1006/jpho.2002.0174
Available online at http://www.idealibrary.com on IDEAL®

# The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian

## Robert McAllister*

*Department of Linguistics, Stockholm University, Stockholm, Sweden*

## James E. Flege

*Department of Rehabilitation Sciences, University of Alabama at Birmingham, Birmingham, U.S.A.*

## Thorsten Piske

*English Department, Kiel University, Kiel, Germany*

*Received 12th May 2000, and accepted 28th January 2002*

This study examined the influence of the L1 phonetic system on the acquisition of Swedish quantity distinctions. The aim was to test the hypothesis that difficulty in acquiring an L2 contrastive category is related to the role in the L1 of the phonetic feature upon which the L2 category is based. Twenty native speakers each of American English, Latin American Spanish and Estonian participated. The phonologies of the subjects' L1 display three degrees of overall prominence of the duration feature. The subjects, who had all lived in Sweden for at least 10 years, were given a production and perception test to assess their mastery of Swedish quantity. The Estonian subjects performed much like Swedish controls. However, some native English, and even more Spanish subjects, differed from native speakers of Swedish. A follow-up experiment compared matched groups of native English and Spanish subjects. It showed that experienced and inexperienced groups of native English and Spanish adults were equally successful in learning to produce and perceive Swedish quantity distinctions, but that the native English subjects more closely matched native Swedish speakers than the native Spanish subjects did. Taken together, our results support the feature hypothesis in that these subjects' success in learning the Swedish quantity contrast seems to be related to the role of the duration feature in the L1. 

*Address correspondence to R. McAllister, Department of Linguistics, Stockholm University, Stockholm, Sweden. E-mails: bob@ling.su.se; jeflege@uab.edu; piske@anglistik.uni-kiel.de

0095–4470/02/$ - see front matter 

## 1. Introduction

A great many recent studies have shown that adults who learn a second language (L2) are apt to produce and perceive L2 phonetic segments differently than do individuals who are monolingual native speakers of the target L2 (e.g., Flege & Fletcher, 1992; McAllister, 1997). Many phonetic dimensions, however, span more than a single phonetic segment. One example of this is the quantity distinctions found in Swedish, which involve a complex interplay between temporal dimensions (i.e., the duration of a vowel and that of the following consonant) and spectral dimensions (i.e., formant values in the vowel). The aim of this study was to provide a better understanding of the role of native language (L1) phonetic features in L2 speech acquisition by examining the production and perception of Swedish quantity distinctions by native speakers of two languages in which quantity distinctions are not phonologically relevant (viz., English and Spanish) and one language in which quantity is an important aspect of the phonology (Estonian).

Numerous studies have supported the importance of L1 transfer. Most of the research done on the transfer issue has been concerned with segmental phenomena (see Bohn (1995) for an overview of vowels and Strange (1998) for an overview of consonants). Transfer at the prosodic level has also been well documented, however, with stress and rhythm receiving the most attention (e.g., Bannert, 1984; Schmid, 1986; Grover *et al*., 1987; Flege & Bohn, 1989; Levitt, 1991; Archibald, 1993). While these studies have contributed significantly to the understanding of the nature of the phenomenon, there are still some important unanswered questions to be addressed. One of these concerns what aspects of the perception and production of the L1 are actually ''transferred''.

The investigation of transfer at the prosodic level such as that undertaken in this study would seem particularly important in light of recent, largely unsubstantiated claims that prosodic errors contribute relatively more to the impression of foreign accent than segmental accuracy (e.g., Pennington & Richards, 1986).

The experiments reported here are designed to test a ''feature'' hypothesis relevant to the learning of L2 phonology. This hypothesis was implied in Flege's speech learning model (SLM) (Flege, 1995), one of the current models of L2 speech acquisition, and states that L2 phonetic category formation may be blocked by a mismatch in the phonetic features used to signal contrast in the L1 and L2. The hypothesis could be explicitly stated as follows: L2 features not used to signal phonological contrast in L1 will be difficult to perceive for the L2 learner and this difficulty will be reflected in the learner's production of the contrast based on this feature.

Speech perception reflects the phonetic properties of the language (or languages) that are heard in infancy and childhood. Even before infants have acquired their first word, they respond to vowel sounds in a language-specific manner (Kuhl, Williams, Lacerda, Stevens & Lindblom, 1992; Kuhl, 1995). Infants appear to learn to attend to features that optimally distinguish contrastive sound units and to ''weight'' these features according to their importance for the L1 phonetic system (Jusczyk, 1992, 1993). This process has come to be referred to as ''perceptual attunement'' (Pisoni & Aslin, 1980). The test of the feature hypothesis proposed here can contribute to the determination of whether adults have the ability to re-attune their speech perception system. More specifically, in this study we investigate adults'

ability to acquire sensitivity to a feature not exploited in the phonetic system of their L1.

There is evidence that the use or weighting of a feature used in the L2 but not in the L1 may present a problem for the L2 learner. Gottfried & Suiter (1997) found that native English adult speakers had little trouble learning vowel quality in Mandarin vowels. However, these learners of Mandarin were less successful in learning lexical tone, a feature not used in English to signal phonological contrast. Viewing these findings in terms of the feature hypothesis stated earlier, they would appear to support the suggestion that an L2 contrastive category will be difficult to acquire if it is based on a phonetic feature not exploited in the L1 to signal phonological contrast. Extended to the present study, it might mean that native speakers of English might be able to learn spectral differences between Swedish vowels, but would have difficulty learning temporally based distinctions in the Swedish quantity contrast. This is because spectral quality, but rarely duration, is used to distinguish word meanings in English. Native speakers of Spanish, where duration is not exploited in the phonology, may have even more difficulty with the duration feature in the Swedish quantity contrast.

As an addition to the list of factors influencing L2 phonological learning, an alternative to the transfer explanation has been offered in some recent work that has emphasized the potential importance of language independent auditorily based strategies in L2 perception. Bohn (Bohn, 1995; see also Bohn & Flege, 1990) studied the perception of American English vowels by native Spanish and German learners. The stimuli were synthetic continua of vowels that varied orthogonally in first formant (F1) frequency and duration. The adult L2 learners tended to rely more heavily on duration when identifying the synthetic English vowels than on spectral quality. Although tense and lax vowels in German are contrasted by means of both durational and spectral cues (Ramers, 1988), one might assume as Bohn does, that native German speakers have a perceptual strategy which includes the use of duration in the perception of vowel distinctions. The finding might be considered surprising, however, for native speakers of Spanish. This is because Spanish does not use duration to signal phonological contrasts, and there are only small phonetic differences in the duration of vowels (Navarro, 1968). On the basis of these and other findings, Bohn formulated a "desensitization hypothesis" which states that if an L2 learner is desensitized to a spectral difference with contrastive function in the L2, durational information will be resorted to in the learner's attempt to hear the contrast, at least in the context of a two-alternative forced-choice test. This hypothesis is based on results which "...suggest that duration cues in vowel perception are easy to access whether or not listeners have had specific linguistic experience with them" (Bohn, 1995, p. 294). If this hypothesis is correct, then, contrary to the feature hypothesis stated earlier, native speakers of English and Spanish should have little difficulty producing and perceiving Swedish quantity distinctions, at least those that are based primarily on duration.

### 1.1. *Quantity in Swedish as an L2*

Traditionally, the primary phonetic feature associated with quantity distinctions in the phonology of a language has been duration differences in the vowels and or consonants, hence the "long–short" or "quantity" terminology. Relatively few

languages exhibit quantity distinctions and there is considerable variation with regard to the use of duration and spectral features in these distinctions. Experimental evidence shows, for example, that in standard French, duration plays only a minor role as a cue in the phonological system (Gottfried & Beddor, 1988; Miller & Grosjean, 1997). Another such language, as mentioned above, is Spanish (Navarro, 1968). This might make it difficult for native French or Spanish speakers to acquire segmental phonetic distinctions in an L2 that are based in whole or part on duration, and *a fortiori*, quantity distinctions.

Among the languages that do signal phonological distinctions by means of differences in duration, there are language-specific relationships between the durational and spectral features in vowels which accompany the realization of the contrast. In English, spectral information in vowels covaries with temporal cues. Although English is traditionally regarded as a language that does not exploit duration in its phonology, studies have shown that native speakers of English are sensitive to the duration feature and are sometimes able to identify a vowel on the basis of its length (e.g., Mermelstein, 1978; Whalen, 1989). The role of duration in contrasts, then, could be viewed as not as prominent in English as in Swedish or Estonian (see the brief descriptions below). However, native speakers of English make somewhat greater use of duration to signal contrast than native speakers of Spanish or French do. Therefore, it might be somewhat easier for speakers of a language like English to learn Swedish quantity distinctions. Thus, we introduce the concept of the prominence of a phonetic feature in the phonology of L1 as having a possible role in the acquisition of the L2 in this study (Gottfried & Beddor, 1988; Miller & Grosjean, 1997). We do not wish to compare these languages in terms of the phonological patterning of duration but only in terms of the degree to which this phonetic feature is used in the contrast patterning and to what extent this degree may or may not facilitate the perceptual re-attunement conducive to the successful acquisition of L2 speech.

In contrast to Spanish where, as noted above, duration seems to be of little or no consequence for the identification of phonological distinctions, Finnish (Lehtonen, 1970) and Estonian (Lehiste, 1997) could be cited as examples of languages in which complex patterns of durational relationships are a salient feature of their phonology. In these languages, the vowel and consonant in a syllable can each have an independent contrastive duration. Unlike Swedish described in the next few paragraphs, long vowels can be followed by either long or short consonants and short vowels can be accompanied by both consonant lengths. Vowel quality is of little or no consequence to the contrasts signaled by temporal cues. Native speakers of a language like Estonian, in which duration plays a prominent role in the quantity contrast, may be aided by this aspect of their L1 in their efforts to learn a second language with a quantity contrast based on duration.

As mentioned earlier, the target L2 in this study is standard Swedish which has a quantity contrast. This aspect of Swedish phonology has been the object of study for over a century (e.g., Aurén, 1874). The most definitive modern description of the quantity system has been presented by Elert (1964).

There are two phonetic aspects of the Swedish contrast which are relevant to the work presented in this paper. The first of these is the distribution of the duration feature. In stressed syllables, there is a complementary relationship between the duration of the vowel and the consonant. When a long vowel occurs, the following

consonant is short. When the vowel is short, the consonant is long, hence the structures V:C and VCC (the CC representing either a long consonant or a cluster). These are the syllable types that will be used as test material in the experiments reported here. This description of the duration feature in Swedish stressed syllables is based on acoustic measurements. Given the regularity of this relationship, it seems likely that the perceptual cues have to do with the durational relation of vowel and consonant although acoustic vowel length alone (or consonant length alone) clearly separates the two phonological classes long and short (Jonasson & McAllister, 1972). No explicit research on the question of the perceptual cues used by native listeners has been done for the duration feature in Swedish but in light of arguments put forward by Diehl & Kluender (1987) which indicate that listeners use all the available acoustic information to perceive speech, it would be reasonable to assume that the duration of both the vowel and the consonant in stressed syllables are perceptually relevant to the realization and recognition of the Swedish quantity contrast.

The second of these relevant phonetic correlates in the Swedish quantity contrast is the relationship between segmental duration and vowel quality. This is a well-documented aspect of quantity in the world's languages (Malmberg, 1944; Durand, 1946, p. 162) and vowel quality differences are associated with quantity in about 20% of the languages which have the contrast (Engstrand, 1986). Several studies have been done to document the relationship between the durational and the spectral features as perceptual cues in this phonological distinction in Swedish. Hadding-Koch & Abramson (1964) found that the primary cue utilized by native Swedish listeners for the recognition of the contrast containing the mid vowel pairs /ɛː/-/ɛ/ and /øː/-/ø/ was duration. Other studies have shown that the syllables containing high and low vowels are distinguished by a combination of duration and vowel quality but that the primary cue used by native listeners in the recognition of the quantity contrast tends to be the spectral features associated with the vowel quality (Engstrand, 1986; Engstrand & Krull, 1994; Behne & Czigler, 1997; Behne, Czigler & Sullivan, 1999). In summary, the phonetic realization of the quantity contrast in Swedish makes this language convenient to use in testing the above-stated feature hypothesis.

The specific aim of this study was to examine the acquisition of the quantity contrast in syllables containing four pairs of Swedish vowels (the mid vowels /øː/-/ø/, /ɛː/-/ɛ/ and the nonmid vowels /ʉː/-/ʉ/, /aː/-/a/). Speakers of three L1s were examined, Latin American Spanish, North American English and Estonian. Our previous discussion of these L1s would allow us to informally categorize these three languages in terms of the exploitation of the duration feature in their respective phonologies. Duration appears to be most prominent in Estonian, and least prominent in Latin American Spanish. In English, where the role of duration is traditionally analyzed as secondary to other features that make up a contrast, it could be argued that its prominence is intermediate between the other L1s in this study. According to the feature hypothesis outlined earlier, the development of new phonetic categories for sounds in an L2 may be blocked by the absence of the contrastive use of a feature in the L1. It would predict that the native speakers of English, in which duration plays a minor role, and Spanish, in which duration seems not to have a phonological function, would have difficulties with the Swedish quantity contrast in syllables with mid vowels where the contrast is based on the

duration feature, whereas the Estonians would be successful in the acquisition of the contrast in these syllables. Because the native English and native Spanish speakers do have spectral differences in the contrasting vowel categories of their L1, the feature hypothesis would predict their having less difficulty learning the Swedish quantity contrast in the high and low vowels, where the contrast is based primarily on spectral differences.

The study is presented in six sections. Section 1 is the Introduction above. Section 2 examines the perceptual sensitivity to Swedish quantity distinctions by 20 native speakers each of Swedish, Estonian, English and Spanish. Section 3 examines the production of the durational aspects of the quantity contrast in Swedish by the same 80 subjects. Section 4 presents an analysis of the individual variation in the results of the preceding two experiments. Section 5 examines the production and perception of Swedish quantity distinctions by native English and Spanish subjects who differ from the subjects in the first two experiments in terms of Swedish-language experience as measured by length of residence in Sweden. One aim of this section is to determine if the differences between native English and Spanish subjects that were observed in Sections 2 and 3 were due to confounded factors. The other aim was to determine if adult learners of Swedish who are relatively experienced in Swedish would perform better than subjects with less Swedish-language experience. Finally, Section 6 is a general discussion of our results in light of our hypothesis.

## 2. Experiment 1

This experiment tested perceptual sensitivity of native speakers of Spanish, English, and Estonian to the quantity distinction between words containing four pairs of phonologically long *vs*. short Swedish vowels: /ʉː/-/ʉ/, /øː/-/ø/, /ɛː/-/ɛ/ and /aː/-/a/. The non-native subjects examined here had all learned Swedish as an L2 in Sweden and all were experienced in Swedish, having lived in Sweden for a minimum of 10 years.

The 80 perceptual stimuli used here were produced by a native Swedish-speaking phonetician. Half of the stimuli were real Swedish words. Of these, half contained a phonologically short vowel (VCC), and half contained a phonologically long vowel (V:C). The remaining 40 stimuli were nonwords created by producing the real words containing a long vowel with the corresponding short vowel and *vice versa*. During the experiment, each word (as well as the paired nonword) was preceded by a phrase that defined it. The task of the non-native subjects, as well as the 20 native Swedish controls, was to determine whether each word and nonword stimulus was a phonologically ''correct'' or ''incorrect'' rendition of the defined word.

The 40 words occur frequently (see Section 2.1.2) in Swedish, and were given a mean familiarity rating of 6.96 (*range:* 6.79–7.00) by the 20 native Swedish-speaking control subjects (on a scale that ranged from 1, ''very unfamiliar'' to 7, ''very familiar''). As will be reported later, the non-native subjects knew the great majority of the words. The lexically oriented task used here, therefore, provided a way to determine if the non-native subjects lexicalized the 40 Swedish words as native Swedish speakers do, that is, with phonologically long (V:C) or short (VCC) vowels as appropriate.

## 2.1. *Method*

### 2.1.1. *Subjects*

Twenty adult native speakers each of Spanish, English, Estonian, and Swedish were recruited in Stockholm, Sweden, and then tested individually in a sound attenuated room at the Department of Linguistics at Stockholm University. The native Spanish speakers were all from Central and South America (mostly Chile and Argentina). The native English speakers were all from North America (mostly the United States). The native Estonian subjects had all been born in Estonia and spoke standard Estonian. All 60 non-native subjects had learned Swedish as an L2 in Sweden, had lived in Sweden for at least 10 years, and reported using Swedish often. The native English and Spanish subjects did not speak any other language having quantity distinctions. As mentioned in the Introduction, quantity distinctions exist in the L1 of the Estonians.

The mean ages of the four groups of subjects are shown in Table I. An ANOVA revealed that the four groups differed in age, $F(3,76) = 26.9$, $p < 0.01$. Tukey's *post hoc* test revealed that the native Swedish subjects were significantly younger than the native Spanish, English and Estonian subjects ($p < 0.05$); and that the native Spanish subjects were younger than the native English subjects ($p < 0.05$).

A series of ANOVAs was carried out to compare characteristics of the three non-native groups (see Table I). These tests revealed that the native Spanish, English and Estonian subjects did not differ significantly in terms of their age of arrival in Sweden, $F(2,57) = 2.8$, $p = 0.07$, self-reported use of Swedish, $F(2,57) = 0.8$, $p > 0.10$, or self-estimated proficiency in Swedish, $F(2,57) = 1.7$, $p > 0.10$. However, the non-native groups did differ in length of residence (LOR) in Sweden, $F(2,57) = 7.9$, $p < 0.01$. Tukey's test revealed that the native English subjects had lived longer in Sweden than the native Spanish and Estonian subjects had ($p < 0.05$), whereas the LORs of the native Spanish and Estonian subjects did not differ significantly.

The possibility that their relatively long LOR in Sweden may have conferred an advantage on the native English subjects will be discussed later in the article. However, with regard to LOR as a predictor of success in L2 phonological acquisition, it should be noted here that the relation between the actual amount of input from Swedish native speakers and our nonnative subjects' LORs was probably related in a nonlinear way. A number of studies have addressed the issue of the

TABLE I. Characteristics of 80 subjects who participated in Experiments 1 and 2

| | Age | *N* | AOA | LOR | L2 use | L2 proficiency |
|---|---|---|---|---|---|---|
| Swedish | 33(7) | 13m, 7f | — | — | — | — |
| Spanish | 43(7) | 12m, 8f | 26(6) | 17(4) | 4.8(1.4) | 5.1(1.0) |
| English | 53(9) | 9m, 11f | 28(7) | 24(8) | 4.9(1.6) | 5.0(1.3) |
| Estonian | 48(8) | 2m, 18f | 31(8) | 17(7) | 5.4(1.3) | 5.6(0.8) |

*Note*: Age, chronological age, in years; AOA, age of arrival in Sweden, in years; LOR, length of residence in Sweden, in years; Swedish use, mean self-rating on a scale ranging from 1 (''very seldom'') to 7 (''very often''); Swedish ability, mean self-rating of ability to read, write, understand, and speak Swedish, on scales ranging from 1 (''very poor'') to 7 (''very good''). Standard deviations are in parentheses.

relevance of LOR. Some of these did not provide evidence that LOR affects L2 performance (e.g., Oyama, 1976; Tahta, Wood & Loewenthal, 1981; Flege, 1988; Thompson, 1991; Moyer, 1999), while other studies have found it to be a good index of performance (e.g., Purcell & Suter, 1980; Riney & Flege, 1998). Two factors seem to affect the results of these studies. One is the LOR ranges that are included. It appears that additional experience with the L2 leads to less foreign-accented speech in the early phases of L2 learning. For highly experienced subjects on the other hand, additional years of experience in the L2 are less likely to lead to a significant decrease in degree of foreign accent. We will have reason to return to this aspect of the role of LOR later in the paper (General Discussion). The other factor is the quality of the input that the learners have received. In a recent reexamination of the role of LOR in L2 performance, Flege & Liu (2000) found that increasing LOR led to a more native-like performance in an L2 for individuals likely to be receiving a large amount of native-speaker input (i.e., students), but not by individuals whose professions (i.e., laboratory researchers) made it less likely for them to receive a substantial amount of native speaker input.

In experiment 3 the difference in LOR mentioned above between the native English- and Spanish-speaking groups will be more closely examined. Other ANOVAs revealed that the three non-native groups did not differ in term of their age upon arrival in Sweden, $F(2,57) = 2.78$, $p = 0.07$, their self-reported use of Swedish, $F(2,57) = 0.8$, $p > 0.10$, or their self-estimated ability in Swedish, $F(2,57) = 1.74$, $p > 0.10$.

A follow-up questionnaire was administered after the study was completed to provide additional information concerning the kind and amount of Swedish input the non-native speakers had received. The information obtained gives a picture of a reasonably homogeneous L2-model. The subjects were asked how and where they learned Swedish. Nearly half (49%) of the 60 non-native subjects were married to or lived with a native Swedish speaker. The breakdown by group (native English, 57%, native Spanish, 30% and native Estonians, 35%) indicated a possible advantage for the native English subjects. Of the 60 non-native subjects, 71% had taken the Swedish courses for immigrants provided by local governments according to Swedish law and consisting of a total of 300–500 h of instruction. The same percentage (71%) of subjects had taken courses in the Swedish adult education program following the national educational curriculum. Most of the subjects (86%) learned Swedish in the Stockholm area and those who did not, learned it in dialect areas close to that of the Mälar valley Central Standard Swedish which exhibit phonological quantity that is realized in the same way as in this standard dialect.

### 2.1.2. *Stimuli*

As already mentioned, 40 naturally produced words and 40 naturally produced nonwords were used as stimuli. As shown in Table II, five words each contained the eight Swedish vowels of interest: /ʉː/, /ʉ/, /aː/, /a/, /øː/, /ø/, /ɛː/, and /ɛ/. Data pertaining to the frequency of the stimuli has been taken from: *Frequency Dictionary of Everyday Swedish* (Allén, 1970) and *Talspråksfrekvenser* (Spoken Language Frequencies) edited by Allwood (1999). The former is composed of one million words from written texts and the latter of 1.2 million words in a transcribed spoken material from various sources. The average frequency of occurrence in these data

TABLE II. The Swedish words that were used as perceptual stimuli in Experiment 1, and that were produced by the subjects in Experiments 2 and 3

| Phonologically short (V/CC) | | | | Phonologically long (V:/C) | | | |
|---|---|---|---|---|---|---|---|
| | Spelling | IPA | Gloss | | Spelling | IPA | Gloss |
| /ʉ/ | *gubbe* | [gθ̞̀bːə] | old man | /ʉː/ | ***stuga*** | [stʉ̀ːga] | cottage |
| /ʉ/ | *kudde* | [kθ̞̀dːg] | pillow | /ʉː/ | *duva* | [dʉ̀ːva] | dove |
| /ʉ/ | ***sucka*** | [sθ̞̀kːa] | to sigh | /ʉː/ | *bjuda* | [bj̞ʉ̀ːda] | to invite |
| /ʉ/ | *sudda* | [sθ̞̀dːa] | to erase | /ʉː/ | ***duka*** | [dʉ̀ːka] | to set the table |
| /ʉ/ | ***tugga*** | [tθ̞̀gːa] | to chew | /ʉː/ | *spruta* | [spɾʉ̀ːta] | to squirt |
| /ø/ | ***mössa*** | [møsːa] | cap | /øː/ | *röka* | [rø̀ːka] | to smoke |
| /ø/ | *gömma* | [j̞ømːa] | to hide | /øː/ | *spöke* | [spø̀ːkə] | ghost |
| /ø/ | ***tömma*** | [tømːa] | to empty | /øː/ | ***böna*** | [bø̀ːna] | bean |
| /ø/ | *rösta* | [røsta] | to vote | /øː/ | ***köpa*** | [ç̞ø̀ːpa] | to buy |
| /ø/ | *trösta* | [tɾø̀sta] | to comfort | /øː/ | *höna* | [hø̀ːna] | hen |
| /ɛ/ | *grädde* | [gɾɛ̀dːə] | cream | /ɛː/ | ***näsa*** | [nɛ̀ːsa] | nose |
| /ɛ/ | ***täcke*** | [tɛ̀kːə] | to cover | /ɛː/ | *hjärna* | [j̞æ̀ːɳa] | brain |
| /ɛ/ | *hjärta* | [j̞æ̀ʈːa] | heart | /ɛː/ | ***städa*** | [stɛ̀ːda] | to clean |
| /ɛ/ | ***snäcka*** | [snɛ̀kːa] | seashell | /ɛː/ | *läsa* | [lɛ̀ːsa] | to read |
| /ɛ/ | *sträcka* | [stɾɛ̀kːa] | to stretch | /ɛː/ | *pärla* | [pæ̀ːɭa] | pearl |
| /a/ | *panna* | [pànːa] | forehead | /aː/ | ***spade*** | [spɑ̀ːdə] | shovel |
| /a/ | ***tacka*** | [tàkːa] | to thank | /aː/ | *mage* | [mɑ̀ːgə] | stomach |
| /a/ | ***stanna*** | [stànːa] | to stop | /aː/ | *skada* | [skɑ̀ːda] | to damage |
| /a/ | *trappa* | [tràpːa] | stairway | /aː/ | ***gata*** | [gɑ̀ːta] | street |
| /a/ | *skratta* | [skɾàtːa] | to laugh | /aː/ | *stava* | [stɑ̀ːva] | to spell |

*Note*: Only the words in boldface were measured acoustically.

bases of the 40 words used as stimuli in this study is 45.1 (S.D. = 72), indicating that they are reasonably common in Swedish. All the items are produced with lexical tone accent 2.

A criterion used in selecting the 40 words was that they not be minimally paired with another word based on quantity. So, for example, ''trappa'' /tràpːa/ (''stairway'') was an acceptable test word because changing its first vowel from short to long and modifying the following consonant in the appropriate way yields ''trapa'' /tràpːa/, which is not a Swedish word. The 40 nonword stimuli were created by changing the phonologically long vowels in the real words to the corresponding phonologically short vowel and *vice versa*. The length of the following consonant was modified, from short to long or long to short, as appropriate.

We expected the Swedish-speaking phonetician to be successful in modifying quantity patterns, for there are jokes, puns and other word games in Swedish that are based on the switching of phonological length. We, nevertheless, analyzed the word and nonword stimuli to confirm that the nonwords had been produced as intended. Both spectral and temporal measurements indicated that the phonetician had made the desired changes.

The frequency of the first and second vowel formants (F1, F2) were measured at the acoustic midpoint of four tokens of each of the eight vowels of interest (/ʉː/, /ʉ/, /øː/, /ø/, /ɛː/, /ɛ/, /aː/, /a/). An equal number of measurements were made

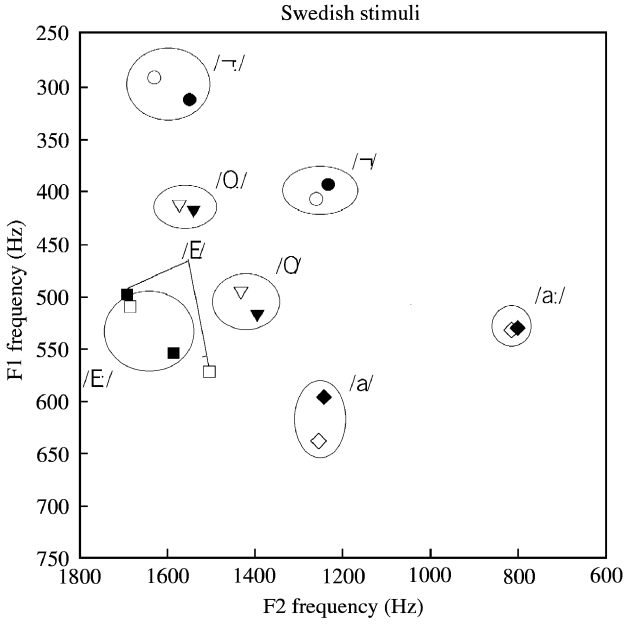


**Figure 1.** The mean frequency of the first and second formants (F1) in Swedish words (filled symbols) and nonwords (unfilled symbols) that contained /ʉː/, /ʉ/, /øː/, /ø/, /ɛː/, /ɛ/, /aː/ or /a/. Each mean is based on four tokens, measured at the acoustic midpoint.

in the corresponding nonwords. The frequency values were obtained at the vowel midpoint of the 64 selected tokens using linear predictive coding analysis. The question of interest was whether the vowels in the words and nonwords were produced with the same spectral quality. For example, would the /ʉː/ tokens produced in words and the /ʉː/ produced in nonwords created by changing the original /ʉ/ to /ʉː/, have similar frequencies? As shown in Fig. 1, the mean frequencies of the eight vowels examined were generally quite similar in the words and nonwords.

We also assessed the temporal characteristics of the words and nonwords. We began by measuring the duration of the vowels and consonants in two words containing each of the eight vowels of interest. (The 16 words measured here were the same words measured in Experiment 2; the criteria used in measurement are described there.) We also measured the corresponding vowel and consonant intervals in the 16 corresponding nonwords. We next computed the ratio of the duration of each measured vowel to the duration of the following consonant. This yielded V:/C ratios for the words containing phonologically long vowels, and V/CC ratios for the words with phonologically short vowels.

We submitted the two ratios described earlier to a (2) Lexical Status $\times$ (2) Phonological Length $\times$ (2) Vowel Type ANOVA to determine if the phonetician produced the words and nonwords with similar quantity differences. This analysis revealed that ratios produced in the words and the nonwords (*means* = 1.32 *vs.* 1.19) did not differ significantly, $F(1,24) = 0.3$, $p > 0.10$, and that the LS factor (word *vs.* nonword) did not interact significantly with any other factor. As expected, the V:/C

ratios were significantly longer than the V/CC ratios (1.87 *vs.* 0.64). The difference between these ratios (*mean* = 1.23) was similar to the difference in ratios (*mean* = 1.26) obtained in Experiment 2 for 20 native Swedish speakers.

Inspection of Fig. 1 also indicates that the phonetician produced somewhat larger spectral differences between phonologically long and short high vowels (/ʉː/-/ʉ/) and low vowels (/aː/-/a/) than between the two pairs of mid vowels (/øː/-/ø/, /ɛː/-/ɛ/). A difference in the extent to which temporal differences were accompanied by spectral differences in mid and nonmid vowels was expected from previous research (Elert, 1964; Hadding-Koch & Abramson, 1964; Jonasson & McAllister, 1972; Engstrand, 1986; Engstrand & Krull, 1994; Behne & Czigler, 1997; Behne *et al.*, 1999).[1]

### 2.1.3. *Procedure*

The word and nonword stimuli described in the last section were randomly presented over headphones a single time to each subject. Each word and the nonword paired to it by changing the phonological length (see above) were presented in separate halves of the experiment. The first author wrote, and then produced and recorded, short phrases that provided an informal definition of each of the 40 words. During the experiment, each word and paired nonword were preceded by the same phrase. For example, the phrase *Ett redskap som används i trädgården för att gräva* ('A tool used in the garden for digging') preceded the real word stimulus *spade* [spɑ̀ːdə ('shovel') and the nonword stimulus **spadde* [spàdːə]. As will be described in Section 3, these phrases were familiar to the subjects. Prior to the perception experiment presented here, the subjects participated in a production experiment, which is presented as Experiment 2. The definition phrases were used to elicit the production of the 40 Swedish words. Their success in producing the desired words confirmed that the phrases were sufficient to define each of the 40 real words.

The subjects were told to listen to each phrase and the stimulus that followed it. Their task was to decide if the stimulus was a correct or incorrect production of the word that was defined. The subjects responded using a mouse. The correct response for the 40 words was "correct", whereas the correct response for the 40 nonwords was "incorrect". Two percent correct scores were computed for each subject, one for the words containing mid vowels (/øː/, /ø/, /ɛː/, /ɛ/) and one for the words containing nonmid vowels (/ʉː/, /ʉ/, /aː/, /a/). Both scores were based on 40 forced-choice judgments. For example, to compute the scores for nonmid vowels, we determined how many of the five real words each containing /aː/, /a/, /ʉː/ and /ʉ/ were labeled "correct", and how many of the five non-words each containing the same four vowels were labeled "incorrect".

## 2.2. *Results and discussion*

Fig. 2 shows the percent correct scores obtained for the words containing mid and nonmid vowels. Two important patterns are evident. First, the native Swedish and

[1]The long member of the /ʉː/-/ʉ/ contrast is usually transcribed as [ʉːβ] to represent the fact that the lips are closed towards the end of vowel (McAllister, Lubber & Carlson, 1974; Bleckert, 1987). The diphthongization in the long vowel might provide another dimension along which /ʉː/ and /ʉ/ might be distinguished perceptually (see, e.g., Hillenbrand & Nearey, 1999).

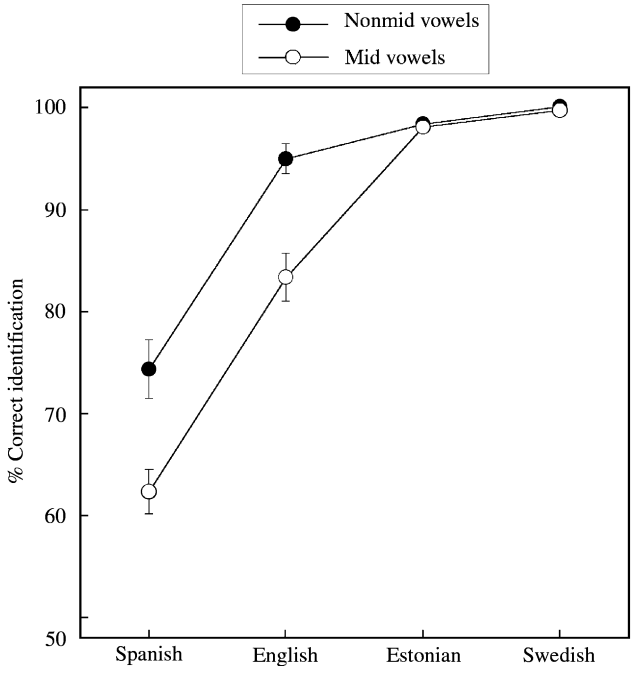


**Figure 2.** The percent correct perception scores obtained for the identification of two pairs of Swedish mid vowels and two pairs of Swedish nonmid vowels by the subjects in four groups.

Estonian subjects responded correctly more often (*means* = 99.8 and 98.1%, respectively) than native English and Spanish subjects did (*means* = 89.2 and 68.4%, respectively). Second, there was little difference in the scores obtained for the words containing mid and nonmid vowels for the native Swedish and Estonian subjects. However, the native English and Spanish subjects obtained higher scores for the words containing nonmid vowels than for those containing mid vowels (native English: 95.0 *vs.* 83.4%; native Spanish: 74.3 *vs.* 62.4%).

The percent correct scores shown in Fig. 2 were submitted to a (4) Native Language × (2) Vowel Type ANOVA. This analysis yielded significant main effects of NL, $F(3,76) = 54.4$, $p < 0.01$, and VT, $F(1,76) = 45.7$, $p < 0.01$, and also a significant NL × VT interaction, $F(3,77) = 14.9$, $p < 0.01$. Test of simple main effects revealed that the interaction had two sources. First, the simple effect of NL was significant for both the mid vowels, $F(3,76) = 63.3$, $p < 0.01$, and the nonmid vowels, $F(3,76) = 31.3$, $p < 0.01$. However, Tukey's tests revealed a different pattern of between-group differences for the mid and nonmid vowels. The native Swedish and Estonian subjects obtained higher scores for the mid vowels than did the native English subjects, whose scores were significantly higher than the native Spanish subjects' scores ($p < 0.05$). For the nonmid vowels, the native Swedish, Estonian and English subjects obtained higher scores than the native Spanish subjects did ($p < 0.05$). However, the native Spanish and English subjects did not differ significantly.

The second source of the NL × VT interaction was a difference between groups for words containing mid and nonmid vowels. The simple effect of VT was significant for the native Spanish, $F(1,19) = 23.6$, $p < 0.01$, and native English subjects, $F(1,19) = 22.3$, $p < 0.01$. However, the differences in the scores obtained for the words containing mid and nonmid vowels was non-significant for the native Swedish, $F(1,19) = 0.3$, $p > 0.10$, and Estonian subjects, $F(1,19) = 0.0$, $p > 0.10$.

In summary, this experiment revealed that the native Spanish, English and Estonian subjects differed in their perception of Swedish quantity distinctions. For words containing mid vowels, the native Swedish and Estonian subjects obtained higher scores than the native English subjects, who in turn obtained higher scores than the native Spanish subjects did. For the nonmid vowels, only the native Spanish subjects differed significantly from the native Swedish subjects. As discussed in the Introduction, the difference between the native English and Spanish subjects might be attributed to a greater importance of temporal distinctions in the phonetic system of English than Spanish. The difference in the non-native subjects' perception of quantity distinctions involving the mid and nonmid vowels may be due to the fact, also discussed in the Introduction, that there are larger spectral differences between the vowels in the pairs /ʉː/-/ʉ/ and /aː/-/a/ than between /øː/-/ø/ and /ɛː/-/ɛ/ (see also Fig. 1). Both the effects of the L1 phonetic system and differences in how quantity distinctions are implemented in the Swedish syllables examined here are discussed at greater length in the General Discussion section.

## 3. Experiment 2

The aim of this experiment was to assess the production of Swedish quantity distinctions by the 20 native and 60 non-native subjects whose perception was assessed in Experiment 1. The subjects produced the 40 Swedish words that were subsequently used as perceptual stimuli in that experiment. As mentioned in the Introduction, tauto-syllabic consonants following phonologically long Swedish vowels are short, whereas consonants following phonologically short Swedish vowels are long. Thus, the dependent variable, examined in this experiment were duration ratios, V/CC ratios for the words containing a phonologically short vowel (/ʉ/, /ø/, /ɛ/, /a/) and V:/C ratios for the words containing a phonologically long vowel (/ʉː/, /øː/, /ɛː/, /aː/).

The native Swedish subjects examined in this experiment produced words having phonologically long vowels with an average V:/C ratio of 2.1, and words having phonologically short vowels with an average V/CC ratio of 0.9. The primary question addressed by this experiment was whether the non-native subjects would produce ratio differences between words containing phonologically short and long vowels that were as large as the differences in ratios produced by the native Swedish speakers. We also wanted to determine if the non-native subjects would be more successful in producing quantity distinctions in words containing nonmid than mid vowels.

### 3.1. *Method*

The subjects produced the 40 Swedish words shown in Table II. They were recorded one at a time using high-quality equipment (Panasonic Model SV-3700 DAT tape recorder, Sennheiser Model MKE2 microphone).

We decided not to elicit the production of Swedish words using written lists because phonological quantity in Swedish words is specified orthographically. Had the subjects read words from a list, they might have produced quantity distinctions in the Swedish words based on the orthography, even if they did not produce quantity distinctions in everyday speech. We therefore used the definition task and showed the subject a clear picture illustrating the word. The first author recorded phrases that informally defined each of the 40 test words. The subjects' task, after hearing each phrase, was to say the target word it defined. For example, "*ett redskap som används i trädgården för att gräva*" ('a tool used in the garden for digging') was used, together with a clear picture of a shovel, to elicit production of the word 'spade' /spàːdə/ ('shovel').

If the subject said the correct target word after hearing its definition (and seeing a picture which illustrated the word), the experimenter nodded. This indicated to the subject that he/she was to repeat the target word at the end of a carrier phrase (*Nu är det* __, 'Now it is __'). If the subject did not say the word—either because it was unknown or because the definition and/or picture was not sufficiently clear—the experimenter provided an L1 translation equivalent. (Once again, a nod from the experimenter indicated that the word should be said at the end of the carrier phrase.) If this did not suffice, the experimenter said the Swedish target word aloud and the subject repeated it at the end of the carrier phrase. As will be described below, few target words had to be prompted in this way. Just as importantly perhaps, hearing a word cued vocally was unlikely to have enabled a subject to produce a quantity distinction if he/she did not normally produce and perceive quantity distinctions in Swedish.

The sentences spoken by each subject were digitized at 10.0 kHz using a personal computer. Temporal measurements of intervals in the test words were later made from digital spectrograms using Cool Edit (Syntrillium Software Corporation). Preliminary analyses revealed that it was difficult to segment some vowels (e.g., those preceded or followed by a glide or liquid) reliably. Therefore, to increase measurement accuracy, and to avoid missing data, we decided to measure just two words containing each of the eight vowels of interest. The words that were selected for acoustic analysis, which are marked in boldface in Table II, were all easily segmented.

We measured the duration of the first vowel in each word and also that of the following consonant. The vowel onset was defined as the point of release of the constriction of the preceding consonant. The acoustic landmark defining this segmentation point varied somewhat across the words that were measured. It was the onset of the release burst for words beginning in a stop (e.g., *duka*), and the point of rapid energy increase in the region of F2 for words beginning in a nasal (e.g., *näsa*) or fricative (e.g., *sucka*). The vowel offset was defined as the point of complete constriction of the following consonant, defined as the end of energy in the region of F2. The ending portion of some vowels were aspirated, that is, contained aperiodic formant structure. These portions were included as part of the vowel interval. The following consonant was measured from the end of the first vowel, as defined earlier, to the release of constriction, as defined for word-initial consonants.

A total of 640 (4 groups $\times$ 20 subjects $\times$ 8 words) V:/C ratios and 640 V/CC ratios were calculated from the vowel and consonant measurements just described. An

average ratio was then calculated based on the two words representing each vowel of interest. This yielded eight ratios per subject, one for each target vowel.

### 3.2. *Results and discussion*

The average ratios obtained for the words containing /øː/, /ø/, /ɛː/, /ɛ/, /ʉː/, /ʉ/, /aː/ and /a/ are shown in Fig. 3. A preliminary analysis revealed that the subjects in all four groups produced significantly different V:/C *vs.* V/CC ratios in words containing phonologically long and short vowels (viz., in /ʉː/-/ʉ/, /øː/-/ø/, /ɛː/-/ɛ/ and /aː/-/a/ $p<0.01$). However, the native Spanish subjects, and to a lesser extent the native English subjects, produced somewhat smaller differences in C:/V *vs.* V/CC ratios than the native Swedish and Estonian subjects did.

We calculated four average ratios from the eight ratios described earlier, two V:/C ratios (one for the words containing the mid vowels /øː/ and /ɛː/, and one for the

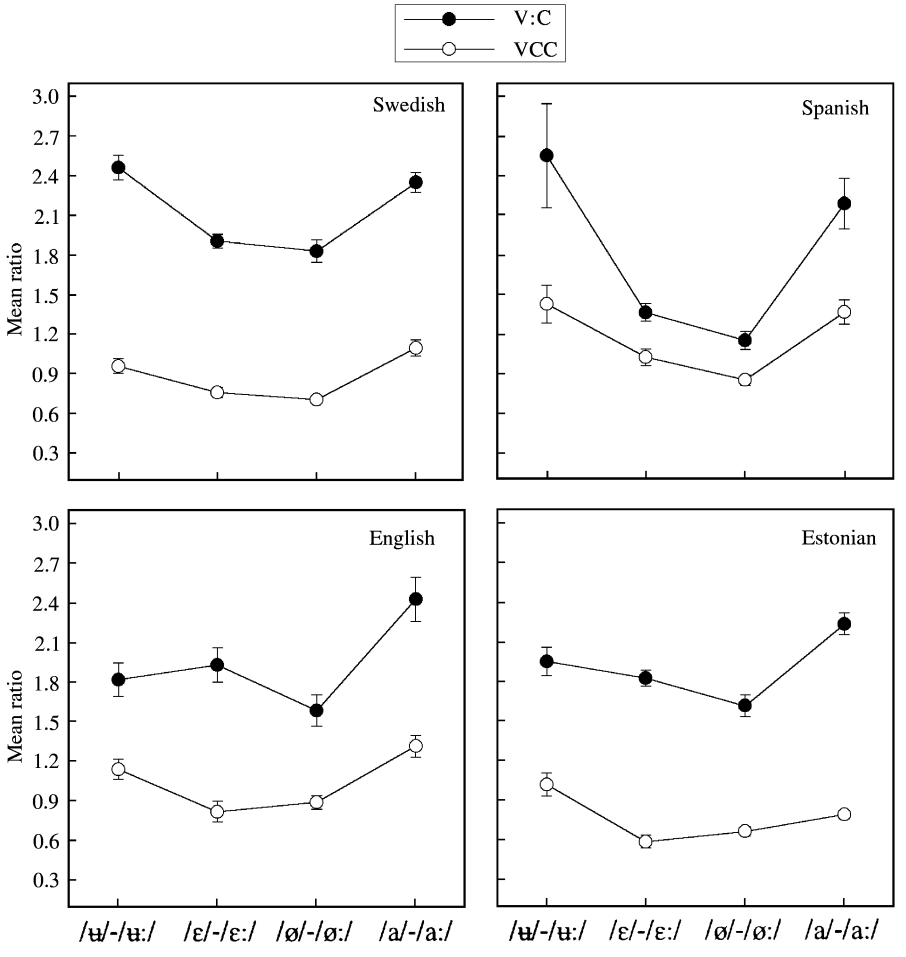


**Figure 3.** The average V:/C *vs.* V/CC ratios produced by the native Swedish, Spanish, English, and Estonian subjects in Swedish words containing phonologically long and short vowels. The error bars enclose ±1 S.E.

words containing the nonmid vowels /ʉː/ and /aː/) and two V/CC ratios (one for the words containing the mid vowels /ø/ and /ɛ/, and one for the words containing the nonmid vowels /ʉ/ and /a/). We then calculated two ratio difference scores from these average ratios. One was the difference in the V:/C and V/CC ratios obtained for the words containing mid vowels (/øː/, /ø/, /ɛː/, /ɛ/). The other ratio difference score was the difference in V:/C and V/CC ratios obtained for the words containing nonmid vowels (/ʉː/, /ʉ/, /aː/, /a/). These difference scores indexed the magnitude of the quantity distinctions produced in words containing mid and nonmid vowels.

The mean ratio difference scores obtained for the four groups of subjects are shown in Fig. 4. The native Spanish subjects produced a smaller ratio difference for mid vowels than the other three groups of subjects did, whereas the four groups differed less for the words containing nonmid vowels. The ratio difference scores shown in Fig. 4 were submitted to a mixed-design (4) Native Language × (4) Vowel Type ANOVA. This analysis yielded significant main effects of NL, $F(3,76) = 10.3$, $p < 0.01$, and VT, $F(1,76) = 14.3$, $p < 0.01$, and also a NL × VT interaction, $F(3,76) = 5.1$, $p < 0.01$.

Tests of simple main effects revealed that a two-way interaction arose because the native Spanish subjects produced smaller ratio differences than the other three groups did, but only for words containing mid vowels. The simple effect of NL proved to be significant for the words containing both mid vowels, $F(3,76) = 25.7$, $p < 0.01$, and nonmid vowels, $F(3,76) = 2.8$, $p < 0.05$. Tukey's tests revealed that the native Spanish subjects produced a smaller difference in V:/C *vs.* V/CC ratios than

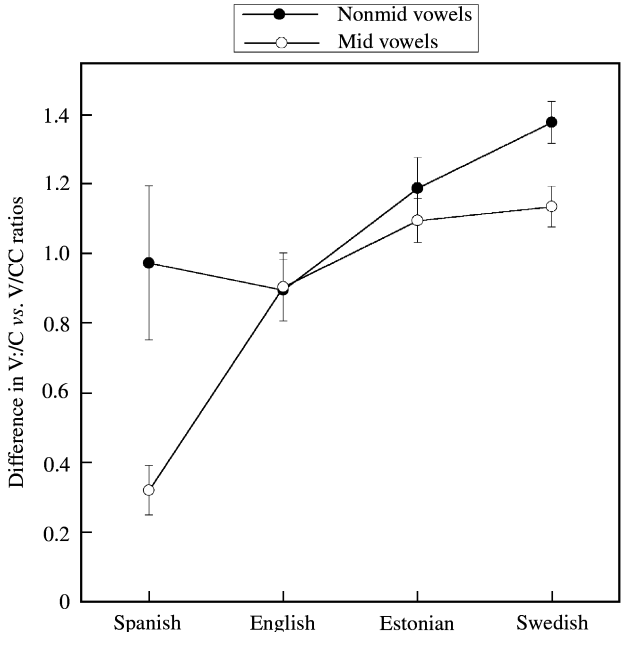


**Figure 4.** The difference in V:/C and V/CC ratios in words containing mid vowels (/øː/-/ø/, /ɛː/-/ɛ/) and nonmid vowels (/ʉː/-/ʉ/, /aː/-/a/) as spoken in Swedish words by native speakers of Swedish, Spanish, English, and Estonian. The error bars enclose ±1 S.E.

the native English, Estonian and Swedish subjects did for the words with mid vowels ($p<0.05$). There were no significant between-group differences for the words containing nonmid vowels, however.

In summary, this experiment yielded fewer significant native *vs.* non-native differences than the experiment examining the perception of quantity distinctions (Experiment 1) did. The native Spanish subjects produced smaller quantity distinctions than the native Swedish subjects did, but only in words containing mid vowels. Recall that in the perception experiment reported in Section 2, both the native Spanish and English subjects obtained significantly lower percent correct scores than the native Swedish subjects did for the words containing mid vowels; and the native Spanish subjects obtained lower scores than the native Swedish subjects did for words with nonmid vowels.

The results of this experiment agree with two tentative conclusions we drew from the results of Experiment 1. The first is that quantity distinctions in mid vowels are more difficult to learn than those in nonmid vowels, perhaps because mid vowel quantity distinctions are based on temporal duration and accompanied by smaller spectral differences than nonmid vowel quantity distinctions are (Hadding-Koch & Abramson, 1964; Engstrand, 1986; Engstrand & Krull, 1994; Behne & Czigler, 1997; Behne *et al*., 1999). The second conclusion is that native speakers of Spanish are less successful in learning Swedish quantity distinctions than native speakers of English are. As discussed in the Introduction, the difference between the native English and Spanish subjects might be attributed to a slightly greater importance of temporal distinctions in the phonetic system of English compared to Spanish.

Finally, it is important to note that the native speakers of Estonian did not differ significantly from the native Swedish subjects in this experiment or in Experiment 1. We attribute this to the fact that quantity distinctions based on duration exist in Estonian. Although Estonian quantity distinctions are implemented differently than those of Swedish (Elert, 1964; Lehiste, 1997), the Estonian subjects' previous experience with the duration feature may have facilitated their learning of this aspect of Swedish phonology. Note that this conclusion implies that, in the language of the Introduction, their attunement to subsegmental "features" transfers from the L1 to an L2.

## 4. Inter-subject variability in perception and production

Experiments 1 and 2 examined the production and perception of Swedish quantity distinctions by native speakers of English, Spanish, and Estonian. As we have pointed out, the Estonian group performed much like the native Swedish group. The native English group, and to an even greater extent the native Spanish group, diverged from the native Swedish group. However, there were considerable inter-subject differences among the native English and Spanish subjects**.** Some native speakers of English and a few native speakers of Spanish performed just like native speakers of Swedish, whereas others performed at a chance level. That is, they were unable to determine if known Swedish words were supposed to have a V:C structure or a VCC structure. The aim of the analyses presented below was to provide insight into the basis for this inter-subject variability. We did so in a series of eight stepwise multiple regression analyses.

This section is organized as follows. We first examined the effect of five "subject" variables, including length of residence in Sweden on the production and perception of contrasts involving mid and nonmid vowels. We then examined the effects of two "lexical variables" on the outcome measures. These two variables were the native Spanish and English subjects' ratings of the familiarity of words containing mid and nonmid vowels, and the number of words with mid and nonmid vowels (*maximum* = 20) that had to be prompted in the production experiment. We reasoned that the need for prompting might indicate lack of familiarity of a word, and that unfamiliarity might have caused perception errors, production errors, or both.

### 4.1. *Subject variables*

The predictor variables used in the four analyses presented here were the native Spanish and English subjects' length of residence (LOR) in Sweden, age of arrival in Sweden, chronological age, average self-reported use of Swedish, and average self-reported use of the L1 (either English or Spanish). The criterion variables examined were: (1) the percentage of correct responses obtained in Experiment 1 for words containing mid vowels; (2) the percentage of correct responses for words containing nonmid vowels; (3) the mean difference in V:/C *vs.* V/CC ratios obtained in Experiment 2 for words containing mid vowels; and (4) the mean difference in ratios for words containing nonmid vowels.

Neither analysis examining the production scores (i.e., ratios) accounted for a significant amount of variance ($p > 0.10$). However, LOR accounted for 11% of the variance in the perceptual scores obtained for words containing mid vowels, $F(1,38) = 4.78$, $p < 0.05$, and 12% of the variance in the perceptual scores obtained for words containing nonmid vowels, $F(1,38) = 5.17$, $p < 0.05$. None of the remaining subject variables accounted for a significant additional amount of variance in the outcome measures.

### 4.2. *Lexical variables*

The predictor variables used in these analyses were the native Spanish and English subjects' familiarity ratings for the 20 words each containing mid and nonmid vowels, and the number of times they had to be prompted on these words during the production experiment.

Following the perception and production tests, the subjects rated the familiarity of the 40 Swedish test words using a scale that ranged from "very unfamiliar" (1) to "very familiar" (7). We calculated the mean rating given by each subject to the 20 words each containing mid and nonmid vowels. The mean ratings were then submitted to a (4) Native Language (NL) $\times$ (2) Vowel Type (VT) ANOVA, which yielded a significant main effect of NL, $F(3,75) = 2.73$, $p < 0.05$, but a non-significant main effect of VT, $F(1,75) = 2.48$, $p > 0.10$, and a non-significant NL $\times$ VT interaction, $F(3,75) = 0.87$, $p > 0.10$. Tukey's test revealed that the native Estonian subjects rated the words as less familiar than the native Swedish subjects (mean ratings of 6.36 *vs.* 6.96, $p < 0.05$), whereas the native Spanish and English subjects' ratings (*means* = 6.51 for both groups) did not differ from the native Swedish subjects' ratings, or differ significantly from one another ($p < 0.05$).

As mentioned earlier (Section 3.1), the non-native subjects were usually able to determine what word to say after hearing a definition (and seeing a picture) of it. The native Swedish subjects never had to be prompted. However, in some cases, aural prompting was sometimes needed to elicit word production by the non-native subjects. We tabulated the number of words with mid and nonmid vowels (*maximum* = 20 each) that each non-native subject had to be prompted on. A (3) Native Language × (2) Vowel Type ANOVA yielded a significant main effect of NL, $F(2,57) = 5.4$, $p < 0.01$, but a non-significant main effect of VT, $F(1,57) = 2.0$, $p > 0.10$, and a non-significant NL × VT interaction, $F(2,57) = 0.0$, n.s. Tukey's test revealed that the Spanish subjects had to be prompted on significantly more words than either the native English or Estonian subjects (*means* = 1.8 *vs.* 0.6 and 0.2, respectively; $p < 0.05$), whereas the native English and Estonian subjects did not differ significantly. Note that the prompting results for the native Spanish subjects contrasts with their familiarity ratings reported in the previous paragraph.

The mean familiarity ratings and the prompting scores just described were regressed onto the production and perception scores obtained for words containing mid and nonmid vowels. The prompting scores accounted for 16% of the variance in the production ratios for the words containing mid vowels, $F(1,38) = 7.0$, $p < 0.05$, and 31% of the variance for the words containing nonmid vowels, $F(1,38) = 16.8$, $p < 0.01$. The prompting scores accounted for 18% of the variance in the percent correct scores obtained for the words having mid vowels, $F(1,38) = 8.1$, $p < 0.01$, and 32% of the variance for the words with nonmid vowels, $F(1,38) = 18.8$, $p < 0.01$. In none of these four analyses did the familiarity ratings account for a significant additional amount of variance in an outcome measure.

### 4.3. *Discussion*

The regression analyses presented here accounted for a modest amount of variance in the native Spanish and English subjects' production and perception of Swedish quantity distinctions. Somewhat more variance was accounted for by what we have termed a "lexical" variable (i.e., the number of words prompted during the production experiment) than by what we have called "subject" variables. The native Spanish subjects had to be prompted on significantly more words than the native English subjects. The prompting scores accounted for 18 and 32% of the variance in the native Spanish and English subjects' perceptual scores for words containing mid and nonmid vowels, respectively; and 16 and 31% of the variance in the production scores obtained for mid and nonmid vowels.

To further consider the role of word knowledge, as indexed by how often words had to be prompted, we examined the data obtained for individual native Spanish subjects. Eight native Spanish subjects did not have to be prompted for any of the words containing mid vowels. The perception scores obtained for these "no-prompt" subjects differed considerably (*range*: 48–85% correct), as did the ratios the no-prompt subjects produced in words containing mid vowels (*range*: −0.20 to 0.88). There were five native Spanish subjects who had to be prompted on three or more words containing mid vowels. The perception scores obtained from these subjects were generally low (*range*: 50–63%), as were their production ratios (*range*: −0.11 to 0.43).

Three native Spanish subjects did not have to be prompted for any of the words containing non-mid vowels. Their perception scores differed considerably (70, 78, 98% correct), whereas the ratios produced by all three no-prompt subjects were substantially smaller (0.57, 0.06, 0.45) than the native Swedish subjects' ratios. Three other native Spanish subjects had to be prompted on three or more words containing nonmid vowels. Their perception scores were uniformly low (43, 53, 65%), but their production ratios differed considerably (0.67, 4.23, 0.16).

Taken together, these data suggest that the subjects who had to be prompted often usually differed considerably from the native Swedish subjects in producing and perceiving Swedish quantity contrasts. One might conclude that their errors derived from a lack of knowledge of particular words. However, such a strong conclusion would probably be unwarranted. First, some of the subjects who never had to be prompted, and thus could be demonstrated to have known all 40 Swedish words, performed poorly in Swedish. Second, even assuming that prompting a word meant it was not known, the prompting data account poorly for the scores. Take, for example, the two native Spanish subjects who were prompted on three or more words containing nonmid vowels, respectively. If these two subjects had responded correctly to all words that they knew (as demonstrated by the absence of prompting), they should have obtained percent correct scores of 85 and 80%, not the scores that were actually obtained (65 and 43%, respectively).

As mentioned in Section 2.1.1, the native Spanish subjects were significantly younger than the native English subjects (*means* = 43 *vs.* 53 years), and had lived for a significantly shorter period of time in Sweden than the native English subjects (*means* = 17 *vs.* 24 years). Of the five subject variables examined in this section, just one—LOR—accounted for a significant amount of variance in the outcome measures (11% of the variance in the perceptual scores for words containing mid vowels, 12% for the words with nonmid vowels). No other subject variable accounted for a significant amount of variance in the production scores.

One possible explanation for differences in the native Spanish and English subjects' perception and production of Swedish quantity distinctions is that duration plays a different role in the Spanish and English phonetic systems (see the Introduction). Another possibility is that some or all of the differences observed in this study were due to differences in the Spanish and English subjects' overall familiarity with Swedish (as indexed by LOR), and their knowledge of the Swedish lexical items (as indexed by the prompting scores). To help choose between these two interpretations, we carried out another experiment comparing groups of Spanish and English subjects that were matched in terms of LOR and number of words that needed to be prompted.

## 5. Experiment 3

The aim of this experiment was to provide greater insight into the differences observed in Experiments 1 and 2 between native speakers of English and Spanish. To do so, we examined the perception and production of Swedish quantity distinctions by groups of native English and Spanish subjects who did not differ in terms of how long they had been speaking Swedish (as indexed by LOR), and in

terms of their knowledge of the Swedish words used as stimuli (as indexed by familiarity ratings and the number of words that had to be prompted to elicit production). We reasoned that if the English subjects outperformed the native Spanish subjects here, as was the case previously in Experiments 1 and 2, it would strengthen the conclusion that differences between the native Spanish and English subjects were likely to have arisen from differences in the Spanish and English phonetic systems not from differences in the subjects' experience with the Swedish language.

### 5.1. *Method*

We recruited 10 native Spanish and 10 native English subjects who were relatively inexperienced in Swedish. As summarized in Table III, the native Spanish and English subjects in the two "short LOR" groups had been living in Sweden for an average of 3.5 and 3.8 years (*range*: 2–5 years), respectively. We also chose pairs of native Spanish and English subjects (10 each) who had participated previously in Experiments 1 and 2. The paired native Spanish and English subjects, who were matched for LOR, had been living in Sweden for an average of 18.1 years (*range*: 10–27 years). The data for these subjects will be presented here for the second time. Thus, a total of 40 subjects, 20 native English (half-inexperienced) and 20 native Spanish (half-inexperienced) were examined here.

To determine if the native Spanish and English groups formed in this way were balanced in terms of LOR, we carried out a Native Language (Spanish *vs.* English) × Experience Level (EL) (short *vs.* long LOR) ANOVA. This ANOVA indicated that the LORs of the native Spanish and English subjects (*means* = 10.8 *vs.* 10.9 years) did not differ significantly, $F(1,36) = 0.01$, $p > 0.10$, and that the NL × EL

TABLE III. Characteristics of the four groups of 10 subjects each who participated in Experiment 3

| | Native Spanish | | Native English | |
|---|---|---|---|---|
| | Short LOR | Long LOR | Short LOR | Long LOR |
| Age | 28.7(4.1) | 45.9(6.4) | 34.7(10.5) | 48.9(8.7) |
| AOA | 25.2(3.9) | 27.7(6.1) | 31.0(10.4) | 30.8(8.5) |
| LOR | 3.5(0.9) | 18.1(3.6) | 3.8(1.4) | 18.1(6.7) |
| L2 use | 4.9(1.9) | 5.2(1.5) | 4.0(2.2) | 4.9(1.4) |
| L2 ability | 4.3(0.6) | 5.0(0.8) | 4.0(1.2) | 4.7(1.4) |
| Fam-NM | 5.5(1.0) | 6.6(0.4) | 4.9(0.6) | 6.4(0.8) |
| Fam-Mid | 5.2(1.1) | 6.6(0.4) | 4.7(0.8) | 6.3(1.0) |
| Prompt-NM | 3.1(2.1) | 1.1(0.7) | 3.1(2.7) | 0.8(1.5) |
| Prompt-Mid | 1.4(1.3) | 1.2(1.2) | 1.7(1.9) | 0.7(1.9) |

*Note*: Age, chronological age, in years; AOA, age of arrival in Sweden, in years; LOR, length of residence in Sweden, in years; L2 use, self-reported use of Swedish; L2 ability, self-reported ability in Swedish; Fam-NM and Fam-Mid, mean familiarity ratings for 20 words each containing nonmid and mid vowels; Prompt-NM and Prompt-Mid, number of words (*maximum* = 20) containing nonmid and mid vowels that had to be prompted during speech elicitation. Standard deviations are in parentheses.

interaction was also non-significant, $F(1,36) = 0.02$, $p > 0.10$.[2] This means that if the native English subjects outperformed the native Spanish subjects examined here, it could not be attributed to a difference in LOR.

We next examined the native Spanish and English subjects' knowledge of the 40 Swedish test words. We calculated how many of the 20 words each containing mid and nonmid vowels each subject had to be prompted on. We also calculated each subjects' average familiarity ratings for the words with mid and nonmid vowels. The prompting scores and familiarity ratings were then examined in separate NL (English, Spanish) × EL (short *vs.* long LOR) × Vowel Type (mid *vs.* nonmid) ANOVAs. The inexperienced subjects had to be prompted on significantly more words than the experienced subjects did (2.3 *vs.* 1.3, $F(1,36) = 6.9$, $p < 0.05$), and they judged the Swedish words to be significantly less familiar than the experienced subjects did (5.1 *vs.* 6.5, $F(1,36) = 33.0$, $p < 0.01$). Importantly, the native Spanish and English subjects did not differ significantly in terms of the number of words prompted (1.7 *vs.* 1.6, $F(1,36) = 0.06$, $p > 0.10$), or in terms of their subjective ratings of word familiarity (6.0 *vs.* 6.5, $F(1,36) = 2.4$, $p > 0.10$). In neither analysis did the NL factor interact significantly with any other factor ($p > 0.10$). This means that if the native English subjects outperformed the native Spanish subjects examined here, it could not be attributed to a difference in lexical knowledge or familiarity.

## 5.2. *Results*

### 5.2.1. *Perception*

The percent correct scores obtained for the words containing mid vowels (/ɛː/, /ɛ/, /øː/, /ø/) and nonmid vowels (/ʉː/, /ʉ/, /aː/, /a/) are shown in Fig. 5. (The 20 Swedish subjects' scores have been juxtaposed as a point of reference, but will not be analyzed here.) Lower scores were obtained for the words containing mid vowels than nonmid vowels. This held true for the native speakers of English and Spanish; and for subjects with relatively short and long LORs in Sweden. The effect of LOR differed as a function of native language. The experienced native English subjects obtained somewhat higher scores than the inexperienced native English subjects did, both for words with mid and nonmid vowels. However, the native Spanish subjects showed a trend in the opposite direction.

The percent correct scores shown in Fig. 5 for the native Spanish and English subjects were submitted to a (2) Native Language × (2) Experience Level × (2) Vowel Type ANOVA. This analysis revealed that the scores obtained by the inexperienced and experienced subjects (*means* = 75 *vs.* 79% correct) did not differ significantly, $F(1,36) = 0.6$, $p > 0.10$. However, the scores obtained for words containing mid and nonmid vowels (71 *vs.* 83% correct) differed significantly, $F(1,36) = 49.6$, $p < 0.01$. The scores obtained by the native Spanish and English subjects (*means* = 69 *vs.* 85% correct) also differed significantly, $F(1,36) = 14.7$, $p < 0.01$. Importantly, the NL factor did not interact significantly with any other factor ($p > 0.10$).

[2]Other two-way ANOVAs revealed that differences between the native Spanish and English subjects for four other variables were non-significant: chronological age, 37.3 *vs.* 41.8, $F(1,36) = 3.3$, $p = 0.08$; the self-ratings of Swedish ability, 4.7 *vs.* 4.3, $F(1,36) = 1.1$, $p > 0.10$; the ages of arrival, 26.5 *vs.* 30.9, $F(1,36) = 3.3$, $p = 0.08$; mean ratings of Swedish use, 5.0 *vs.* 4.5, $F(1,36) = 1.1$, $p > 0.10$.

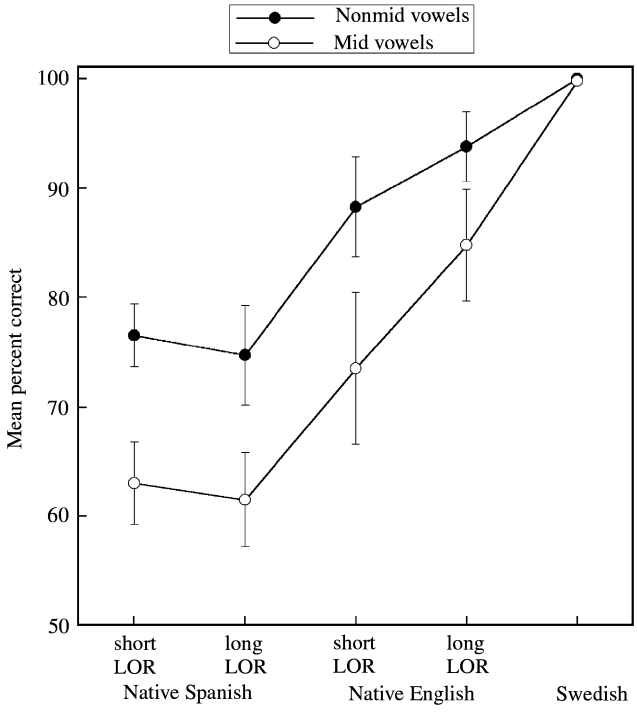


**Figure 5.** The percent correct scores obtained in a perception experiment examining words containing mid and nonmid vowels. The four groups of subjects differed in native language and length of residence in Sweden. The data obtained for 20 native Swedish subjects have been juxtaposed to the results obtained in Experiment 3. The error bars enclose $\pm 1$ S.E.

### 5.2.2. *Production*

We next examined the 40 subjects' production of Swedish vowels. As in Experiment 2, we computed the ratio of each vowel of interest to that of the following consonant. We then computed the difference in the V:/C and V/CC ratios obtained for the words containing mid vowels (/ɛː/-/ɛ/, /øː/-/ø/) and nonmid vowels (/ʉː/-/ʉ/, /aː/-/a/).

The ratio difference scores obtained for the four groups (along with those obtained for the native Swedish subjects) are shown in Fig. 6. The experienced and inexperienced native Spanish subjects, and the inexperienced native English subjects, obtained lower scores for words containing mid than nonmid vowels. However, a trend in the opposite direction was obtained for the experienced native English subjects. Once again, the effect of LOR seemed to differ as a function of native language. The experienced native English subjects obtained somewhat higher scores than the inexperienced native English subjects did, both for words with mid and nonmid vowels. However, the native Spanish subjects showed a trend in the opposite direction.

The ratios shown in Fig. 6 were examined in a (2) Native Language $\times$ (2) Experience Level $\times$ (2) Vowel Type ANOVA. The ratios produced by the

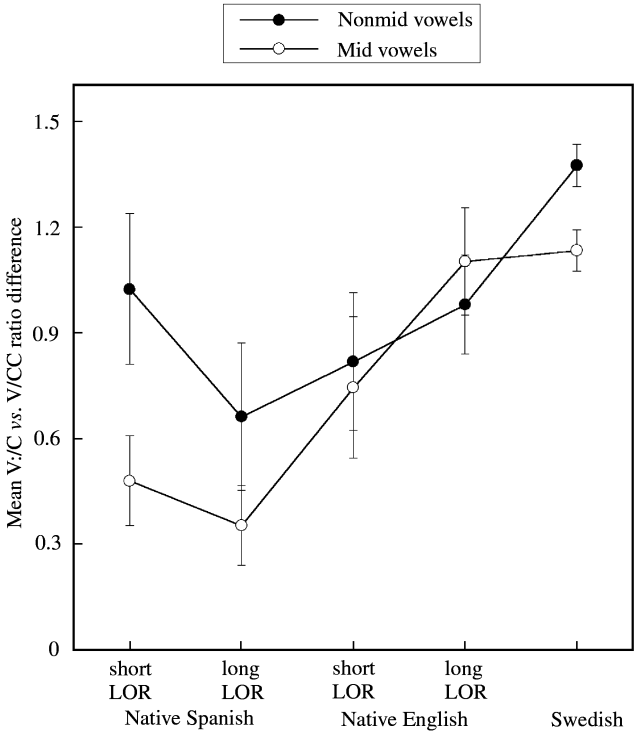


**Figure 6.** The difference in V:/C and V/CC ratios in words with mid and non-mid vowels spoken by four groups of subjects differing in native language and length of residence in Sweden. The data obtained for 20 native Swedish subjects have been juxtaposed. The error bars enclose $\pm 1$ S.E.

inexperienced and experienced subjects (*means* = 0.78 *vs.* 0.77) did not differ significantly, $F(1,36) = 0.003$, $p > 0.10$. However, the ratios obtained for the words containing mid and nonmid vowels (0.67 *vs.* 0.87) did differ significantly, $F(1,36) = 6.0$, $p < 0.05$.

Of greatest interest was the effect of native language. The NL factor was only marginally significant, $F(1,36) = 3.4$, $p = 0.07$. It did not interact significantly with EL, $F(1,36) = 2.7$, $p > 0.10$, or enter into a significant three-way interaction, $F(1,36) = 0.01$, $p > 0.10$. However, a significant NL $\times$ VT interaction was obtained, $F(1,36) = 7.6$, $p < 0.01$.

Simple effects tests revealed that the NL $\times$ VT interaction had two sources. First, the native English subjects produced significantly larger—and thus more Swedish-like—ratios in the words containing mid vowels than the native Spanish subjects did (0.92 *vs.* 0.42), $F(1,38) = 10.8$, $p < 0.01$. However, the native English and Spanish subjects' ratios for words containing nonmid vowels (0.90 *vs.* 0.84) did not differ significantly, $F(1,38) = 0.09$, $p > 0.10$. Second, the effect of Vowel Type differed for the native Spanish and English subjects. The native Spanish subjects produced larger ratio differences in the words containing nonmid than mid vowels (0.84 *vs.* 0.42), $F(1,19) = 8.9$, $p < 0.01$. However, the native English subjects' ratios for nonmid and mid vowels (0.90 *vs.* 0.92) did not differ significantly, $F(1,19) = 0.09$, $p > 0.10$.

### 5.3. *Discussion*

The question addressed here was whether native speakers of English are better able to learn Swedish quantity distinctions than native speakers of Spanish. Unlike the groups examined in Experiments 1 and 2, the native Spanish and English subjects examined here did not differ significantly in LOR or in terms of the number of words that needed to be prompted to elicit word production. The native English subjects obtained significantly higher perception scores than the native Spanish subjects (*means* = 85 *vs.* 69%). The effect of native language was only marginally significant in the production experiment ($p = 0.07$). This is because the native English subjects produced significantly larger, more Swedish-like, ratios in the words containing mid vowels but not in the words containing nonmid vowels. These results support the view that the native English subjects were able to learn Swedish quantity distinctions more successfully than the native Spanish subjects because of the phonetic structure of their native language, rather than due to differences in lexical knowledge or Swedish-language experience.

## 6. General discussion

The primary aim of this study was to determine if the phonetic structure of the L1 would affect the learning of Swedish quantity distinctions. Experiment 1 examined the perception of Swedish quantity distinctions by 20 native speakers each of Spanish, English and Estonian who had learned Swedish as an L2 in adulthood. Experiment 2 examined the same subjects' production of quantity distinctions. Both experiments focused on words containing pairs of mid vowels (/øː/-/ø/, /ɛː/-/ɛ/) and nonmid vowels (/ʉː/-/ʉ/,/aː/-/a/) differing in phonological quantity. As discussed in the Introduction, previous research has established that duration is the phonetic feature used to realize the quantity distinction in Swedish mid vowels whereas two potential perceptual cues, vowel quality as well as duration, accompany the quantity distinction in Swedish nonmid vowels (Elert, 1964; Hadding-Koch & Abramson, 1964; Jonasson & McAllister, 1972; Engstrand, 1986; Engstrand & Krull, 1994; Behne & Czigler, 1997; Behne *et al.*, 1999). This led us to formulate a "feature" hypothesis. According to this hypothesis, L2 features not used to signal phonological contrast in L1 will be difficult to perceive for the L2 learner and this difficulty will be reflected in the learner's production of the contrast based on this feature.

In Experiment 1, the native Swedish controls and the Estonian subjects obtained higher perceptual scores for words containing mid vowels than the native English subjects, who obtained higher scores than the native Spanish subjects. For words with nonmid vowels, the native Spanish subjects, but not the native English (or Estonian) subjects, differed significantly from the native Swedish subjects. The dependent variable examined in Experiment 2 was the size of differences in V:/C and V/CC ratios that subjects produced in words containing phonologically long and short vowels. Only the native Spanish subjects were found to have produced significantly smaller ratio differences in quantity distinctions than the native Swedish subjects, and then only in words containing mid vowels.

The results of Experiments 1 and 2 suggested that the native Estonian subjects were more successful in learning Swedish quantity contrasts than the native English subjects, and that the native English subjects were more successful than the native Spanish subjects. As mentioned in the Introduction, one possible explanation for the difference between the Estonian subjects, on the one hand, and the native Spanish and English subjects, on the other, is that Estonian, unlike Spanish and English, has quantity distinctions based on the duration feature as a prominent aspect of its phonology. Although Estonian and Swedish quantity distinctions differ in terms of their phonetic and phonological structures (Elert, 1964; Lehiste, 1997) the Estonian subjects' attunement to the duration feature used in the quantity distinctions in their L1 appears to have facilitated their learning of Swedish quantity distinctions.

The differences we observed between the native Spanish and English subjects, on the other hand, may be due to the fact that duration plays a less prominent role in distinguishing phonetic segments in Spanish than in English. Recall that native speakers of English are sensitive to the duration feature and are sometimes able to identify a vowel on the basis of its length (e.g., Mermelstein, 1978; Whalen, 1989) but that in Spanish duration has no phonological function. These results, viewed in light of work by Gottfried & Beddor (1988) and Miller & Grosjean (1997), would seem to reflect the prominence of the duration feature in the L1. Future research on a refinement of the feature hypothesis in the formulation of a "feature prominence" hypothesis would appear to be motivated. According to this hypothesis, the relative importance of a feature in the L1 will determine the extent to which the feature is successfully used in producing and perceiving phonological contrasts in the L2.

There were reasons for caution in accepting the conclusion that the native English subjects outperformed the native Spanish subjects because duration plays a more important phonetic role in English than in Spanish. The native Spanish and English subjects examined in Experiments 1 and 2 differed in LOR, and also in how many words had to be prompted to elicit production, which suggested a difference in knowledge of the Swedish word stimuli. Regression analyses revealed that LOR accounted for a small but significant amount of variance in the perceptual scores obtained for the native Spanish and English subjects in Experiment 1 both for words containing mid and nonmid vowels. Therefore, an alternate explanation for the observed Spanish–English differences was that the Spanish and English subjects differed in Swedish-language experience. The regression analyses also revealed that the number of words prompted accounted for a small but significant amount of the variance in the quantity distinctions that the native Spanish and English subjects produced in words containing mid and nonmid vowels.

The aim of Experiment 3 was to choose between the two competing explanations for why the native Spanish subjects were less successful than the native English subjects in learning Swedish quantity distinctions. We compared the performance of groups of native English and Spanish subjects who were matched for LOR, and who did not differ in terms of their knowledge of the Swedish words used as stimuli (as indexed by familiarity ratings and the number of words that had to be prompted to elicit production). The native English subjects obtained significantly higher perceptual scores than the native Spanish subjects. This held true both for words with mid and nonmid vowels. The effect of Native Language was only marginally significant in the analysis of word production. This is because the native English subjects produced larger differences in V:/C and V/CC ratios than the native

Spanish subjects did for words containing mid vowels (/ɛː/-/ɛ/, /øː/-/ø/) but not for the words containing nonmid vowels (/ʉː/-/ʉ/, /aː/-/a/). These results supported the view that the Spanish and English subjects differed because of a difference in the role of duration in the Spanish and English phonetic systems, not because the two groups differed in Swedish-language experience.

As discussed in the Introduction, Bohn (1995) offered a "desensitization" hypothesis as an alternative to the familiar concept of L1 transfer. According to this hypothesis "...duration cues in vowel perception are easy to access whether or not listeners have had specific linguistic experience with them" (Bohn, 1995, p. 294). It is not unreasonable to assume that basic language-independent auditory mechanisms would play an important role in the perception of the sounds of an L2. This might mean, for example, that all adult learners of Swedish as an L2, regardless of their L1 background, would be able to perceptually distinguish words differing in quantity because of the substantial vowel and consonant duration differences that distinguish words with phonologically long and short vowels. We have, in our study, however, not found support for the desensitization hypothesis. Some of the native English subjects examined in Experiment 1, and an even larger number of native Spanish subjects, seemed unaware of whether known Swedish words contained phonologically long or short vowels. This might be interpreted to mean that despite the availability of durational information, the native Spanish speakers did not differentiate words added to their Swedish lexicon based on quantity distinctions.

An interesting detail in this regard is the production and perception performance of the inexperienced native Spanish learners (Figs. 5 and 6). They were, as a group, slightly more successful than their experienced native Spanish-speaking counterparts in Experiment 3. Although this difference was not statistically significant, the slightly higher group average scores might be interpreted along the lines of Bohn's hypothesis in that some of the individuals in this group could have been using the duration information that was available whereas the experienced learners, for some reason, were not able to do the same. Another possible, if speculative interpretation of these results might be that they support findings of LOR studies mentioned in Section 2.1.1. As we pointed out, these studies (e.g. Oyama, 1976; Tahta *et al.*, 1981; Flege, 1988; Thompson, 1991; Moyer, 1999), indicate that additional experience with the L2 leads to less foreign-accented speech in the early phases of L2 learning. For highly experienced subjects, on the other hand, additional years of experience in the L2 are less likely to lead to a significant decrease in degree of foreign accent. These results appear to support an assumption that if the Swedish quantity distinction is learned, it is learned early, i.e. in the first 3–5 years of contact to the L2. Yet another speculative interpretation of these results could suggest that the motivational variable could be behind the lack of a significant performance difference between the long LOR and the short LOR groups. L2 learners may be particularly motivated in the early phases of L2 learning whereas learners who have already been exposed to the L2 for a long time may, on the other hand, have lost some of their initial motivation. These speculative interpretations are, of course, problematic in that they do not seem to be relevant to the performance of the native English subjects in this experiment.

The present study leaves unanswered a number of important questions that must be addressed in future research examining the production and perception of phonetic distinctions in an L2. One question is whether the native Spanish and English

subjects who obtained low scores in the perception experiment would be able to distinguish Swedish words containing long and short vowels in a task that did not involve word recognition, for example, a discrimination task. It is possible that a non-native subject who does not know if a particular Swedish word contains a phonologically long or short vowel might be able to distinguish minimally paired Swedish words differing in quantity (e.g., *mäta* "to measure" *vs. mätta* "satisfied"). A related question is whether the non-native subjects who obtained high scores in the perception experiment presented in this study, and thus closely resembled the native Swedish subjects, perceived Swedish quantity distinctions exactly like the native Swedish subjects. Behne *et al.* (1999) carried out a parametric mani-pulation of duration and spectral (F1, F2) cues to Swedish quantity distinctions. They found that native English learners of Swedish differed from native Swedish speakers in giving greater weight to duration differences than the native Swedish subject did. It would be interesting to determine if, in an experiment like that of Behne *et al.*, native Spanish and English learners of Swedish would be found to differ in terms of the weight they accorded spectral differences (see also Flege, Bohn & Jang, 1997).

Our study confirms the common observation that some adults are able to re-attune their L1 perceptual systems and achieve native-like production and perception of L2 speech. The experimental results indicate that re-attunement of the L1 phonological system to new features, in this case duration, can be difficult for those L2 learners whose L1 does not exploit the feature in question. The native Estonians were most successful in learning the Swedish quantity contrast, the native Spanish speakers were the least successful and the native English speakers' performance was better than the native Spanish speakers but worse than the native Estonians. Therefore, we might conclude that the performance of these L1 groups varies as a function of the prominence of the role of duration in the L1. We have thus found support for the feature hypothesis stated in the Introduction and may have found the basis for a refinement of this hypothesis to a "feature prominence hypothesis" which would relate an L2 learner's success to the role of a given phonetic feature in the L1. The problem of an explicit specification of the concept of "phonological prominence" of a feature is, however, not a trivial one and provides a starting point for future research. Some individuals, however, were able to master the instances of the Swedish quantity contrast based on duration despite the lack of phonological function for the duration feature in their L1. This is not surprising if one accepts that success in learning L2 speech is due to several factors which influence individuals in diverse ways.

The research was funded by a grant from the National Institute for Deafness and Other Communicative Disorders (DC00257). We thank Upali Lockridge for her able assistance and Olle Engstrand for his help as a native speaker of Swedish, and Katsura Aoyama for comments on an earlier version.

## References

Allén, S. (1970) *Frequency dictionary of everyday Swedish*. Stockholm: Almqvist & Wiksell.

Allwood, J. (1999) Talspråksfrekvenser. *Gothenburg Papers in Theoretical Linguistics*, *Gothenburg University, Department of Linguistics*.

Archibald, J. (1993) The learnability of English metrical parameters by adult Spanish speakers, *International Review of Applied Linguistics*, **31,** 129–142.

Aurén, J. A. (1874) *Bidrag till svenska språkets quantitetslära*. Linköping.

Bannert, R. (1984) Problems in learning Swedish pronunciation and in understanding foreign accent, *Folia Linguistica*, **18,** 193–222.
Behne, D. & Czigler, P. (1997) Swedish quantity and quality: a traditional issue revisited, *Reports from the Department of Phonetics, Umeå University*, **4,** 81–83.
Behne, D., Czigler, P. & Sullivan, H. (1999) Perceived vowel quantity in Swedish: native and British listeners, *Journal of the Acoustical Society of America*, **195,** 1095(A).
Bleckert, L. (1987) *Centralsvensk diftongering som satsfonetisk problem. (Diphthongization in Central Sweden as a Problem of Sentence Phonetics)*. Skrifter utgivna av institutionen för nordiska språk vid Uppsala Universitet.
Bohn, O.-S. (1995) Cross language speech perception in adults: first language transfer doesn't tell it all. In *Speech perception and linguistic experience: issues in cross language research* (W. Strange, editor), pp. 279–304. Timonium, MD: York Press.
Bohn, O.-S. & Flege J. E. (1990) Interlingual identification and the role of foreign language experience in L2 vowel perception, *Applied Psycholinguistics*, **11,** 303–328.
Diehl, R. & Kluender, K. (1987) On the categorization of speech sounds. In *Categorical perception: the groundwork of cognition* (S. Harnad, editor), pp. 226–253. Cambridge: Cambridge University Press.
Durand, M. (1946) Voyelles longues et voyelles brèves: essai sur la nature de la quantité vocalique. Collection linguistique. Societé Linguistique Paris, XLIX, Paris: Klincksieck.
Elert, C.-C. (1964) *Phonologic Studies of Quantity in Swedish*. Uppsala: Monografier utgivna av Stockholms kommunalförvalting 27.
Engstrand, O. (1986) Durational correlates of quantity and sentence stress: a cross-language study of Swedish, Finnish and Czech, *Working Papers in Phonetics*, *Phonetics Laboratory, Department of Linguistics, University of California, Los Angeles*, **63,** 1–25.
Engstrand, O. & Krull, D. (1994) Durational correlates of quantity in Swedish, Finnish and Estonian: cross-language evidence for a theory of adaptive dispersion, *Phonetica*, **51,** 80–91.
Flege, J. (1988) Factors affecting degree of perceived foreign accent in English sentences, *Journal of the Acoustical Society of America*, **84,** 162–177.
Flege, J. E. (1995) Second language speech learning: theory, findings and problems. In *Speech perception and linguistic experience: theoretical and methodological issues* (W. Strange, editor), pp. 233–277. Timonium, MD: York Press.
Flege, J. E. & Bohn, O.-S. (1989) An instrumental study of vowel reduction and stress placement in Spanish accented English, *Studies in Second Language Acquisition*, **11,** 35–62.
Flege, J. E., Bohn, O.-S. & Jang, S. (1997) The effect of experience on nonnative subjects' production and perception of English vowels, *Journal of Phonetics*, **25,** 437–470.
Flege, J. E. & Fletcher, K. (1992) Talker and listener effects on the perception of degree of foreign accent, *Journal of the Acoustical Society of America*, **91,** 370–389.
Flege, J. E. & Liu, S. (2000) The effect of experience on adults' acquisition of a second language, *Studies in Second Language Acquisition* **23,** 527–555.
Gottfried, T. & Beddor, P. (1988) Perception of temporal and spectral information in French vowels, *Language and Speech*, **31,** 57–75.
Gottfried, T. L. & Suiter, T. L. (1997) Effect of linguistic experience on the identification of Mandarin tones, *Journal of Phonetics*, **25,** 207–231.
Grover, C., Jamieson, D. & Dobrovolsky, M. (1987) Intonation in English, French and German: perception and production, *Language and Speech*, **30,** 277–295.
Hadding-Koch, K. & Abramson, A. (1964) Duration versus spectrum in Swedish vowels: some perceptual experiments, *Studia Linguistica*, **XVIII,** 94–107.
Hillenbrand, J. & Nearey, T. (1999) Identification of resynthesized /hVd/ utterances: effects of formant contour, *Journal of the Acoustical Society of America*, **105,** 3509–3523.
Jonasson, J. & McAllister R. (1972) Foreign accent and timing: an instrumental phonetic study, *Papers from the Institute of Linguistics, University of Stockholm* (*PILUS*), **14,** 11–40.
Jusczyk, P. (1992) Developing phonological categories from the speech signal. In *Phonological development, models, research, implications* (C. Ferguson, L. Menn & C. Stoel-Gammon, editors), pp. 17–64. Timonium, MD: York Press.
Jusczyk, P. (1993) From general to language-specific capacities: the WRAPSA Model of how speech perception develops, *Journal of Phonetics*, **21,** 3–28.
Kuhl, P. K. (1995) Mechanisms of development and change in speech perception and language. In *Proceedings of the XIIth international congress of phonetic sciences* (K. Elenius & P. Branderud, editors), Vol. 2, pp. 132–139. Stockholm: KTH and Stockholm University.
Kuhl, P., Williams, K., Lacerda, F., Stevens, K. & Lindblom, B. (1992) Linguistic experience alters phonetic perception in infants by 6 months of age, *Science*, **255,** 606–608.
Lehiste, I. (1997) The search for phonetic correlates in Estonian prosody. In *Estonian prosody: papers from a symposium* (I. Lehiste & J. Ross, editors), pp. 11–33. Tallinn: Institute of Estonian Language and Authors.

Lehtonen, J. (1970) *Aspects of quantity in standard Finnish*. Studia Philologica Jyväskyläensia VI. Juväskylä: Gummérus.
Levitt, A. (1991) Reiterant speech as a test of non-native speakers' mastery of the timing of French, *Journal of the Acoustical Society of America*, **90,** 3008–3018.
Malmberg, B. (1944) Die Quantität als phonetisch-phonologischer Begriff. Lunds Universitets Årsskrift, avd. I, bd 41 nr 2. Lund.
McAllister, R. (1997) Perception and production of a second language and the concept of foreign accent. In *New Sounds 97* (J. Leather & A. James, editors), pp. 206–214. University of Klagenfurt.
McAllister, R., Lubker, J. F. & Carlson, J. (1974) An EMG study of some characteristics of the Swedish rounded vowels, *Journal of Phonetics*, **2,** 267–278.
Mermelstein, P. (1978) On the relationship between vowel and consonant identification when cued by the same acoustic information, *Perception and Psychophysics*, **23,** 331–336.
Miller, J. L. & Grosjean, F. (1997) Dialect effects in vowel perception: the role of temporal information in French, *Language and Speech*, **40,** 277–288.
Moyer, A. (1999) Ultimate attainment in L2 phonology: the critical factors of age, motivation, and instruction, *Studies in Second Language Acquisition*, **21,** 81–108.
Navarro, T. (1968) *Studies in Spanish phonology*. Coral Gables: University of Miami Press.
Oyama, S. (1976) A sensitive period for the acquisition of a non-native phonological system, *Journal of Psycholinguistic Research*, **5,** 261–283.
Pennington, M. & Richards, J. (1986) Pronunciation revisited, *TESOL Quarterly*, **20,** 207–225.
Pisoni, D. B. & Aslin, R. (1980) Some developmental processes in speech perception. In *Child phonology, Vol. 2, Perception* (G. Yeni-Komshian, J. Kavanagh & C. Ferguson, editors). New York: Academic Press.
Purcell, E. & Suter, R. (1980) Predictors of pronunciation accuracy: a reexamination, *Language Learning*, **30,** 271–287.
Ramers, K. (1988) *Vokalqualität und -quantität im Deutschen*. Tübingen: Niemeyer.
Riney, T. J. & Flege, J. E. (1998) Changes over time in global foreign accent and liquid identifiability and accuracy, *Studies in Second Language Acquisition*, **20,** 213–244.
Schmid, B. (1986) A comparative study of children's and adults' acquisition of tone accents in Swedish, *Language Learning*, **36,** 185–210.
Strange, W. (1998) Perceptual assimilation of American English vowels by Japanese listeners, *Journal of Phonetics*, **26,** 311–344.
Tahta, S., Wood, M. & Loewenthal, K. (1981) Foreign accents: factors relating to transfer of accent from the first language to a second language, *Language & Speech*, **24,** 265–272.
Thompson, I. (1991) Foreign accents revisited: the English pronunciation of Russian immigrants, *Language Learning*, **41,** 177–204.
Whalen, D. (1989) Vowel and consonant judgements are not independent when cued by the same information, *Perception and Psychophysics*, **46,** 284–292.