The Acquisition of Phonological Structure: Distinguishing Contrastive from Non-Contrastive Variation

Christiane Dietrich

Cover design & illustrations: Norma Varley
Printed and bound by Ponsen & Looijen bv, Wageningen
ISBN: 90-76203-24-5

THE ACQUISITION OF PHONOLOGICAL STRUCTURE: DISTINGUISHING CONTRASTIVE FROM NON-CONTRASTIVE VARIATION

Een wetenschappelijke proeve
op het gebied van de
Sociale Wetenschappen

Proefschrift

te verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen,
op gezag van de Rector Magnificus prof. dr. C.W.P.M. Blom,
volgens besluit van het College van Decanen
in het openbaar te verdedigen
op dinsdag 23 mei 2006
des ochtends om 10.30 uur precies.

door
Christiane Dietrich
geboren op 25 april 1974
te Aken, duitsland.

Promotor:	Prof. dr. A. Cutler
Copromotor:	Dr. D. Swingley (MPI en University of Pennsylvania)
Manuscriptcommissie:	Prof. dr. L. Verhoeven Prof. dr. W. Klein Prof. Dr. A. Christophe (CNRS te Parijs)

The research reported in this thesis was supported by a grant from the Max-Planck-Gesellschaft zur Förderung der Wissenschaften, München, Germany.

For my grandmother.

ACKNOWLEDGEMENTS

Where shall I begin? First and foremost, I would like to thank Anne Cutler, my promoter and Dan Swingley, my day-to-day supervisor for providing me with this invaluable opportunity! Thank you for your expertise, your patience and for continuing to believe that I was still going to finish my dissertation despite a number of obstacles along the way. Thank you also to the members of my reading committee. Thank you to ALL my past and present colleagues in the Comprehension Group for your comments, suggestions and support. Thank you Rian, for the many things you helped me with along the way (such as the binding and posting of my dissertation to the Leescommittee).

November 1st, 1999 was my official starting day at the MPI and Anne Pier's was shortly thereafter. We shared an office space for two of the three years I spent at the MPI. I could not have asked for a better person to share an office with. Thank you Anne Pier for your calming nature, your technical assistance, your encouragement and, most importantly, your friendship. Thank you also for agreeing to translate the summary of my dissertation into Dutch!

Among the many colleagues and friends in the Netherlands whose presence made all the difference to my life and ability to focus while working on my dissertation were Anne Pier, Anita, Aoju, Birgit, Claudia, Elizabeth, Keren, Mandana and Mirjam.

Danke Anita, dass du meine Squashpartnerin warst und für deine nette Gesellschaft beim Reisbreiesssen. Thank you Aoju, for our many walks and conversations and for letting me stay at your place the night before I moved to Vancouver. Danke Birgit, Claudia and Mandana, für eure Freundschaft, Ermunterung und Unterstützung und für viele nette Besuche. Danke auch für die zahlreichen 'Lifts' nach Kranenburg und Kleve (Mandana und Claudia). Claudia, ich vermisse noch immer unsere Nachmittagsspaziergänge. Thank you Elizabeth for your 'open' office door and for providing valuable comments on an early version of my dissertation. Just so everybody knows: If you ever hear footsteps on the second floor of the MPI at 3 a.m., there is no need to worry about burglars – it is likely Elizabeth still at work! Dank je Mirjam, voor goede gesprekken, gezellige avonden en voor het lekker eten! Dank je ook voor de 'lift' naar een bushalte in Groesbeek toen ik na mijn bezoek in Nijmegen vorig jaar mijn bus naar Kleef had gemist. Thank you Aoju, Keren and Mirjam for your timely help answering my many questions about the promotie.

Vielen Dank auch an Frank, der während meines Besuchs in Kanada mein Untermieter war, und mir immer sofort alle Post hinterher geschickt hat.

I would not have been able to complete my research without the help of the MPI Technical Group and many other people at the MPI! In het bezonder, dank je John (Nagengast) – zonder jou had ik nooit mijn experimente op kon zetten. Dank je Ad voor je technische steun en voor het 'open houden' van mijn 'account' (toen ik er al officieel niet meer was) und danke Gerd für deine technische Hilfe und Unterstützung. Thank you Angela (Khadar), for your help running studies and scheduling babies. Verder wil ik Agnes, Angela (Heuts), Uschi, José, Hans, Jan en Pim danken

In Canada, thank you Janet (Werker) for having been an inspirational collaborator, advisor and source of support throughout a great part of my doctoral studies and beyond. Thank you for believing in me and for offering me a postdoc even though I had not completely finished the editing of my dissertation. I wish to thank my friends and former colleagues in Vancouver: Laura, Laurie, Nenagh and Ramesh. Thank you Alisa, Carla, Erin, Marissa and Vashti for helping with the running of the Headturn Study as well as with scheduling of babies.

Thank you to Norma Varley for designing the cover of my dissertation. Een bijzonder woord van dank ook voor Irene Salverda en Paul, mijn paranimfen!

Most importantly, it is hard to know where life would have taken me without the support, encouragement and patience of Paul (once again!) and my family. Although my grandmother will not be there in person during my defense, I know she will be there in spirit.

TABLE OF CONTENTS

1 INTRODUCTION

In speech perception some acoustic variation is relevant to the phonology of the language and some is not. In English, a difference in length of voice onset time (VOT) for voiced and unvoiced stops is a cue that is relevant. Yet vowel duration information is not, at least not with respect to the vowel itself. This is to say native English listeners do not use the duration of a vowel as a primary cue to its identity. Hence lengthening of a vowel does not change it to another vowel.

However, there are languages that do distinguish vowels by length (such as Japanese) and others that do not distinguish voiced from unvoiced stops by VOT (such as Mandarin Chinese) or that make more distinctions along the voicing continuum than English (such as Thai; Maddieson, 1984).

One of many challenges for infants learning English is that they cannot simply learn to ignore acoustic variation of vowel length. This is because vowel duration is important for determining the voicing status of postvocalic plosives (e.g., 'kit' versus 'kid'). Hence vowel length needs to be attended to for some aspects of processing but ignored for others. Indeed, there is evidence suggesting that even newborns can discriminate very small differences in vowel duration of approximately 15 ms (Christophe, Dupoux, Bertoncini & Mehler, 1993).

This dissertation explores infants' learning about phonemically relevant and irrelevant variation in speech perception. The test case used here is vowel length. The approach taken is cross-linguistic comparison of two populations. In one, vowel duration is phonemic (i.e. meaningful for telling words apart in the ambient language), in the other it is not.

The research reported here concerns infants learning about phonology (i.e. the way sounds are organized in the native language). This learning starts early in life – before infants acquire the meanings of words, let alone produce their first words. Infants do learn sound structure, however, and this learning lays the foundation for their acquisition of the phonology of their language. Learning about sound structure is challenging because the realization of sounds is affected by many factors, including the precise phonetic context in which each sound occurs.

There are two important stages during the first year. From birth infants have the capacity to discriminate many of the world's speech sounds, and they maintain this ability for six to eight months, thus allowing them to acquire any languages around them relatively effortlessly. By the end of the first year of life infants come to ignore certain contrasts that are irrelevant in the ambient language.

The precise relevance of these findings for language learning remains unclear. Many early studies looked at infants' discrimination of sound contrasts that do not occur in the ambient language at all. It may not be surprising that infants come to ignore those types of sound contrasts that they are never normally exposed to. Studies of this nature are obviously crucial for looking at language-specific attunement (i.e. increasingly heightened sensitivity to those sound contrasts that are meaningful in the native language). Yet, they leave open a crucial question, namely how infants deal with variability that is in the speech signal but that is nonetheless irrelevant in the language around them in the special sense of being a primary cue for telling sounds and therefore words apart. We will explore this question experimentally and cross-linguistically with respect to vowel length.

To set the present work in context, a brief account of infants' auditory perception will be provided first, since hearing obviously affects infants' processing of the incoming speech stream. This is followed by a discussion of what challenges infants have to deal with as they learn to break up the incoming speech stream of the ambient language into word forms. Among these challenges are learning to compensate for contextual variability in the speech signal and breaking up connected speech into possible word forms. Finally, the role of vowel length in different languages will be discussed, with particular emphasis on English and Dutch.

1.1 Listening in infancy

Audition is precocious relative to vision. Although infants have a preference for face-like configurations, this preference appears to be more general in nature in the sense that it remains when faces are partly jumbled (Maurer & Young, 1983) and there is no specific preference for the mother's face. But there is such a preference for the mother's voice (DeCasper & Fifer, 1980). Audition is already well developed at birth. Infants have even been demonstrated to recognize and prefer a story that had been read to them by their mother in the last three months of pregnancy when that same story was read to them by a stranger after birth (DeCasper & Spence 1986). This indicates that infants are already starting to take in information about certain aspects of the language of the environment before birth. Indeed, many aspects of infants' hearing are effectively mature by the time they are about six months old (Werner, 1996; Werner & Gray, 1998), although three-month-olds still require sounds to be at increased amplitude to be able to discriminate them as well as adults and in some respects the auditory system continues to mature until children are approximately ten years old. This early maturation of the auditory system clearly contributes to – or paves the way for – infants' phonemic discrimination abilities.

1.1.1 Phonetic category formation - attunement to native language phonemic structure

Research on speech perception has revealed that normally developing infants' perceptual sensitivities are extremely sophisticated shortly after birth. Newborns are capable of perceiving phonemic contrasts in all of the world's languages that have been tested to date shortly after birth (Eimas, Siqueland, Jusczyk & Vigorito, 1971; for a review see Aslin, Jusczyk & Pisoni, 1998). This is to say, infants' perceptual sensitivities appear not to have been influenced as yet by the phonemic structure of their native language. It is across the first year that selective fine-tuning of speech-perceptual skills takes place – the 'native-language attunement' referred to above.

For consonants, the influence of the native language on infants' speech perception has been demonstrated from the end of the first year: ten-month-olds have already learnt to pay special attention to those acoustic features that differentiate the phonemes of the ambient language. However, they no longer discriminate most non-native phonetic contrasts (Pegg & Werker, 1997; Werker & Tees, 1984), even if they readily discriminated the same contrasts and even the same tokens at six months. For vowels, native language attunement takes place as early as six months (Kuhl, Williams, Lacerda, Stevens & Lindblom, 1992; Polka & Werker, 1994).

There are some exceptions to these findings, however: some non-native consonants are still easily discriminated, not only by infants but also by adults. These exceptions include Zulu clicks (Best, McRoberts & Sithole, 1988), which do not overlap with any of the native language contrasts – at least for English speakers – and they may therefore be resistant to perceptual decline (Best, 1994). Hence there appear to be differences among non-native sounds in terms of the ease with which they are discriminated by non-native listeners (for a review, see Werker & Lalonde, 1988).

Interestingly, differences have also been found between contrasts that have phonological status and those that do not: Werker and Lalonde (1988) tested six- to eight-month-olds and eleven- to thirteen-month-old English-learning infants on a synthesized place of articulation continuum from /ba/ to /da/ that was either phonemic in both English and Hindi (labial vs. dental or alveolar), only in Hindi (dental versus retroflex) or in neither language (arbitrarily chosen near /da/ end of continuum and likely not phonemic in any of world's languages). As predicted, and in line with previous research, the older infants showed a perceptual decline when discriminating the Hindi-only contrast whereas there was no change in discrimination ability of the English contrast. Interestingly, even the younger infants had difficulty with the arbitrary non-phonemic 'place of articulation' contrast that is not meaningful in either English or Hindi. This suggests that both phonemic relevance and linguistic status of contrasts may have played a role here.

Taken together, the early findings suggest that infants' perceptual sensitivities enable them to discover the pertinent phonetic categories used in their native language across the first year. The following section will explain why this is not a trivial task.

1.2 Problems that need to be tackled by the language-learning infant

1.2.1 Variability problem

Speech perceptual cues are highly variable and can hence not be relied upon in isolation. Although listeners are sensitive to information such as transition time and use this cue as a primary cue when discriminating or categorizing syllables such as /ba/ and /wa/, reliance on this cue in isolation is insufficient because it is not stable. Transition duration information, for example, is affected by differences in speaking rate. Nonetheless, infants, like adults, are able to correctly categorize phonemes despite these rate of speech differences (Miller & Eimas, 1994). For example, whereas the category boundary for a /ba/ and /wa/ stimulus continuum falls between 16 and 40 msec when subjects are presented with syllables that are 80 msec in length, this boundary falls between 40 and 64 msec when stimulus length is increased to 296 msec (Miller & Lieberman, 1979; cited in Miller & Eimas, 1994). Infants hence appear to treat transition duration in relation to later occurring information, namely overall syllable length. This suggests that sensitivity to rate of speech and compensation for its impact on syllable duration are either innate, or acquired very early in life.

A further example reflecting the importance of context-sensitive processing is coarticulation. Anticipatory and perseverative coarticulation and assimilation are among the processes which can 'naturally' alter the status of phonemes in the speech stream of native speakers of English: For example, the nasal in the phrase /in the/ is generally fronted (i.e. realized at a more dental place of articulation) because it is followed by a dental fricative. Assimilation can even result in replacement of entire phonemes, as in /hot pot/, which is sometimes realized as [hop‿pot] in running speech. The presence of these natural processes requires that the young listener be sensitive to these cues for word recognition to be successful. Variation of allophones, the sex and the age of the speaker complicate this featural and contextual variation.

In short, the speech stream lacks invariant acoustic information both at the phoneme level and inevitably also at the word form level. This is problematic since it has obvious implications for infants' ability to break up the incoming speech stream reliably into phonemically well-specified word forms. Nevertheless, infants have been found to be extremely skilled at this from an early age.

Given the substantial variability that potential acoustic cues are subject to, it is clear that infants ultimately need to learn to use multiple cues simultaneously and to ignore certain types of variability when faced with the task of segmenting the speech stream into possible word forms (see section 1.2.4 for more detail).

1.2.2 Segmentation problem

While infants are discovering the phonemic distinctions of their language, they also have to discover words and hence they have to learn to isolate word forms from continuous speech. This is a nontrivial problem because of the lack of invariance of speech sounds and word forms described above. To complicate matters further, there are no clear boundaries between words. Instead, words run into each other in continuous speech such that they cannot be identified simply by segmenting the speech stream into a sequence of familiar and likely word forms based on reliable criteria (e.g. pauses between words). Could word forms initially be learnt in isolation? This is unlikely since only nine percent of utterances addressed to infants appear to be monosyllabic words (van de Weijer, 1999), hence providing only limited opportunity for vocabulary learning. It is noteworthy that van de Weijer's count excluded vocatives, fillers and greetings. The exclusion of these several types of monosyllabic utterances may be argued to explain why the percentage of one-word utterances was so low. Yet, the excluded utterance types do not seem to be words that would provide considerable assistance in the segmentation of continuous speech in that they do not reveal important information regarding the content of a message or with regard to distributional cues, for example.

In cases where word forms are presented in isolation, this certainly provides an additional learning opportunity. However, it is likely that infants do not learn the majority of early word forms in isolation before recognizing them in continuous speech, although some researchers have claimed that infants' first words are indeed those that are likely to be heard in isolation (e.g. Brent & Siskind, 2001). Woodward & Aslin found that even when mothers were instructed to teach their children new words, only 20% of these words were presented in isolation (Woodward & Aslin, 1990). This finding corroborates the hypothesis that isolated words are the exception rather than the rule in speech addressed to young infants, and that infants therefore have to develop means to recognise word forms in context at an early stage. Recent support for the notion that isolated words are not essential for infants to learn to recognize novel words quickly and efficiently was obtained by Kooijman, Hagoort and Cutler (2005) and is discussed in more detail in section 2.1.

Among the other segmental cues that infants appear to be sensitive to are statistical, distributional, phonotactic, allophonic and coarticulation cues (Saffran, Aslin & Newport, 1996a; Saffran, Newport & Aslin, 1996b; Mattys & Juszcyk, 2001; Jusczyk, Hohne & Bauman, 1999; also see 1.2.4).

Models of spoken word recognition by adults suggest that recognition of words involves continuous activation and subsequent competition of potential lexical candidates as more featural, phonological and prosodic information becomes available; competition between candidate words is thought to continue until disambiguating information becomes available (for a review of adult models of word recognition, see Frauenfelder & Floccia, 1998; Norris, McQueen & Cutler, 2000; McQueen, 2005). However, infants cannot rely on their lexicons for implicit segmentation of the speech signal or to disambiguate between numerous lexical candidates, because they do not know enough words to parse the sentences they hear. It is hence clear that the nature of their processing must differ from that of adults – although adults listening to a foreign language are likely to be in a similarly challenging situation. Despite this similarity, there is an important difference between adult second language learners and infants' segmentation: Adults know implicitly many of the regularities present in their native language and use these regularities to arrive at the most likely parse given the signal. Among these regularities are the relative predominance of trochaic (strong-weak) words versus iambic (weak-strong) words in English. Adult listeners become sensitized to their native language rhythm and acquire the appropriate segmentation strategy. Interestingly, it is that same strategy that they appear to resort to when trying to break up a foreign language into possible word forms; this seems to be the case, even when it is not the most useful strategy for the type of language and irrespective of their proficiency in that language (Cutler, 2000). Even bilinguals seem to be capable of employing one strategy only, namely the one required by their dominant or their preferred language (Cutler, 2000). Given that even as yet unborn 'infants' are able to perceive rhythmic aspects of speech (Abrams, Gerhardt, Huang, Peters & Langford, 2000) it is perhaps not surprising that the acquisition of language-specific segmentation skills may have their origin in early childhood. Indeed, infants have been shown to be sensitive to suprasegmental cues, particularly in the early stages. As will be discussed in the following section, infants' early sensitivity to the rhythmic structure of their native language can greatly assist them both in locating possible word boundaries and in processing larger units such as phrases and, more generally, in identifying their native language shortly after birth.

It has recently been suggested that infants not only combine a number of different sources of information to segment potential words from continuous speech in the initial process of speech perception, but that they also favour the use of specific types of cues at different stages of development even though they may still be capable of using the 'dormant' cues under certain task conditions. For example, although infants have been shown to be sensitive to and make use of statistical cues to aid their word segmentation by eight months of age (e.g. Saffran et al., 1996a; Johnson & Jusczyk, 2001), when statistical cues are pitted against stress eight-month-olds give greater weight to stress (Johnson & Jusczyk, 2001).

1.2.3 Sensitivity to rhythmic and prosodic information

Newborns have some basic sensitivity to rhythmic types used in the world's languages, as well as being sensitive to certain aspects of the prosodic structure of their native language from as early as two months (DeCasper & Fifer, 1980). This appears to allow infants to distinguish languages from different rhythmic classes such as stress-timed English from syllable-timed French or mora-timed Japanese (Jusczyk, Cutler & Redanz, 1993; Nazzi, Bertoncini & Mehler, 1998).

Newborns have been shown to be able to discriminate their native language from a foreign language as well as two foreign languages from each other (e.g. Mehler, Jusczyk, Lambertz, et al., 1988; Nazzi, Bertoncini & Mehler, 1998, using the High Amplitude Sucking (HAS) Paradigm).

Slightly older infants have also been tested using this methodology. In one such study, Mehler and colleagues tested American two-month-olds on their ability to distinguish American English and Italian (syllable-timed) prosody (Mehler, et al., 1988). It was expected that infants would have no difficulty distinguishing the two languages. This was indeed found. However, as the language materials used contained both segmental and prosodic information, Mehler and colleagues decided to isolate the prosodic cues by low-pass filtering the stimuli and thus removing most segmental cues. Infants were found to continue to be able to discriminate the two languages under these more controlled conditions. Interestingly, unlike newborns, two-month-olds were unable to discriminate low-pass filtered French and Russian utterances, both of which differ rhythmically from their ambient language.

Given this difference in the processing of two foreign languages between newborns and two-month-olds, Christophe and Morton (1998) set out to address the question of how detailed newborns' representations of their native language may be: they tested newborns' processing of two prosodically similar languages (i.e. English and Dutch) by contrasting English newborns' processing of Dutch both with English and with Japanese.

Two different outcomes were considered: first English newborns may perceive Dutch as 'English' and hence as 'native' (since both languages are stress-timed), in which case they should not discriminate the two. If so, English newborns would be expected to discriminate Dutch from Japanese though (since Japanese is rhythmically different from Dutch/English). The second outcome considered possible was that English newborns would perceive Dutch as different from English and hence as 'non-native' (since stimuli had not been low-pass filtered and hence additional cues were present to aid infants in their discrimination). In the latter case the English newborns may not discriminate Dutch from Japanese just as two-month-olds did not discriminate two foreign languages from each other (Mehler, et al., 1988), as described above. In Christophe and Morton's study, newborns were found not to be homogeneous in their discrimination performance and evidence for both hypothesized outcomes was observed. According to Christophe and Morton (1998), one possible explanation for their findings is that infants were at a transition stage: some had already started noticing language-specific differences between Dutch and English and others had not. If infants were indeed at a transition stage, a longitudinal study of one-to four-month-olds should show a distinct change in discrimination performance such that one-month-olds may treat Dutch as native and four-month-olds may treat Dutch as foreign. Only the older group would be expected to be able to discriminate Dutch from English, whereas only younger group would be expected to discriminate Dutch from Japanese.

These findings are consistent with the conclusion that two-month-old infants are already sensitive to the prosodic structure of their native language, which allows them to discriminate their native language from those foreign languages that differ rhythmically from it. However, they are not as yet able to discriminate non-native languages when both these languages differ rhythmically from the ambient language irrespective of whether both languages differ rhythmically from each other. Newborns' discrimination performance differs from that of two-month-olds in that they are still able to discriminate two foreign languages.

It should be noted, however, that some vowel information may still have been available to the two-month-old infants after the speech had been low-pass filtered and hence it may not have been exclusively prosody that infants' discrimination was based on in the studies described above. Ramus (2002) questioned the above conclusions about the role of prosody in early infant language discrimination on the basis that 'prosody does not reduce to rhythm'. He considered it possible that intonation may play a role in the observed discriminations as well.

How may prosody help the infant to locate possible word-boundaries in the evolving speech stream? Given that 90% of English content words in speech follow a strong-weak stress pattern (Cutler & Carter, 1987), it could be beneficial to assume that strong syllables mark the occurrence of a word boundary (i.e. the onset of a new word). Cutler and Norris have referred to this process as the Metrical Segmentation Strategy (MSS; Cutler & Norris, 1988; Cutler, 1990).

In an investigation of the potential role that infants' sensitivity to the predominant stress pattern of their native language may play in language development, Jusczyk et al. (1993) found that American nine-month-olds, unlike six-month-olds, showed significant preferences for the strong-weak stress pattern of their native language over a weak-strong pattern when tested with the Headturn-Preference paradigm. This preference again remained when the stimuli were low-pass filtered, indicating that it is unlikely to have been the segmental information that infants detected in these stimuli. These findings are in line with the conclusion that infants are becoming increasingly attuned to the characteristic stress pattern of the ambient language between six and nine months. Moreover, the results raise the possibility that American nine-month-olds make use of the MSS (see Swingley (2005a), for an account of how English learners discover this parsing preference).

This is supported by a study by Jusczyk, Houston and Newsome (1999), which revealed that seven-and-a-half- month-old infants familiarized with isolated words having a strong-weak syllable pattern were subsequently able to extract these words from six sentence passages. Extraction of these words turned out to be unsuccessful when the passage only contained the strong syllables (e.g. 'king' from 'kingdom'). The latter finding is interesting in that it confirms that infants had not just responded to the stressed syllable in the previous experiment. Hence their representations of these words must have contained more detail than merely the strong syllable. In a further experiment infants failed to detect words having a weak-strong syllable pattern in passages. At ten-and-a-half months, however, infants were more successful at segmenting weak-strong words from the same text passages. The older infants must either have used a different strategy altogether or they may have complemented the MSS with another more advanced strategy that allowed them to detect words that follow a weak-strong stress pattern. Candidate strategies may be based on phonotactic or allophonic information. It is also conceivable that simply being familiar with a greater number of words may have enhanced the older children's performance at segmenting the weak-strong words in that knowing at least some words that have a weak-strong pattern may have made the infants more aware of the fact that such words do exist. Even if the weak-strong words themselves were unfamiliar to the infants, the fact that infants were most likely able to identify a greater number of the surrounding words at ten-and-a-half months might have assisted them in postulating more correct word-boundaries. In some cases weak-strong words may remain as 'residue' and hence may eventually be recognized as a potential word form.

A more recent study by Houston, Santelmann and Jusczyk (2004) indicated that seven-and-a-half-month-olds' word segmentation abilities even allow them to extract trisyllabic (strong-weak-strong) word forms from fluent speech, as long as the initial syllable carries the primary stress (e.g. cantaloupe). In contrast, when the final syllable carried the primary stress (e.g. cavalier), infants were unable to recognize the entire word; instead they only extracted the stressed syllable (i.e. –lier). Houston and colleagues' findings thus replicate the abovementioned findings by Jusczyk and Aslin (1995), corroborating infants' use of stress cues during segmentation of words from continuous speech as well as extending Jusczyk and Aslin's findings to trisyllabic words.

In addition to showing sensitivity to linguistic prosody, young children also seem to show sensitivity to familiar pieces of music despite changes in key (Trehub, Thorpe, & Morongiello, 1985; in Jusczyk, 1997). This finding relates to infants' ability to compensate for different sources of variability noted above. It also further supports the idea that young children have some ability to process and subsequently recognise certain frequent prosodic patterns.

Taken together, the findings discussed so far in this section suggest that infants have at least some sensitivity to the prosodic structure of their native language from two months of age, which allows them to discriminate their native language from non-native languages, providing the latter differ rhythmically from the ambient language. The languages tested to date include English, Japanese, French, Dutch, Italian and Norwegian (see Jusczyk, 1997, for a review. The segmentation of the speech stream based on prosodic information may also enable infants to detect semantic units in a sentence, insofar as tone groups tend to correspond to semantic units (Ladefoged, 1993). The former may or may not be syntactic units.

Summing up, it has been shown that early sensitivity to prosodic structure and the rhythm of one's native language can fulfill a number of functions. First, it seems to allow infants from at least some language communities to differentiate rhythmically distinct languages and to attend more selectively to their native language and thus learn about its relevant and meaningful cues. Secondly, it enables infants to begin to 'bootstrap' the ambient speech stream according to prosodic groupings. This is crucial given that the majority of speech addressed to young infants tends not to consist of isolated words. The exact role of the Metrical Segmentation Strategy proposed by Cutler and Norris (1988) in infant speech segmentation may become clearer as more data from non-English infants become available.

The studies discussed do not provide any information about infants' relative or differential attention to prosody when other potentially useful sources of information are present simultaneously. This relates to cue integration and hence the process of weighing up different simultaneously present potential cues, an issue which will be discussed further below.

1.2.4 Using multiple cues

Given the substantial variability that potential acoustic cues are subject to, infants need some means to use multiple cues simultaneously when postulating word boundaries while processing the incoming speech stream. The different cues may then be compared in terms of their reliability and compatibility in a specific context. Research on how and when young children develop the skills to integrate several cues, and how they weight two or more cues in relation to each other, has been emerging.

In a study of infants' integration of segmental and suprasegmental cues, Morgan and Saffran (1995) compared infants' likelihood of grouping a pair of syllables while segmental and suprasegmental cues were varied in a systematic way. There were four conditions: in Condition 1 both segmental and suprasegmental regularities supported grouping. In Conditions 2 and 3 either segmental or rhythmic regularities supported syllable grouping. In Condition 4 there were no regularities supporting grouping of syllables. Nine-month-old infants – unlike six-month-olds were found to group the pair of syllables only in Condition 1, where segmental and rhythmic cues were compatible. For six-month-olds rhythmic regularities seemed to predominate and hence no evidence was provided that two regularities (i.e. the presence of segmental and suprasegmental cues) were advantageous. These findings were interpreted as reflecting nine-month-old infants' ability to integrate segmental and suprasegmental information to represent a pair of individual syllables as a single higher-order unit. Such integration of cues may help infants to identify potential words more reliably.

Integration of phonetic and prosodic cues has also been shown to assist infants to process and subsequently recall phonetic information more accurately (Mandel, Jusczyk & Kemler Nelson, 1994). Perhaps infants' recall was enhanced because cue integration allowed them to store and subsequently retrieve a more complete representation, including context-sensitive detail. It is interesting that infants' ability to integrate these two cues closely precedes their language-specific processing of consonants. It may be easier from a processing point of view for infants to focus their attention on language-specific features of consonants once their percept is more stable. The automatic integration of prosodic features with segmental information may hence free up resources thus allowing infants to focus their attention on something new (e.g. language-specific aspects of consonants).

Mattys, Jusczyk, Luce and Morgan (1999) explored nine-month-old infants' use of phonotactic versus prosodic cues to word segmentation using the Headturn Preference Paradigm. The stimulus materials consisted of bisyllabic CVC-CVC nonwords whose cross-syllabic CC-clusters varied in terms of likelihood of occurring at a syllable-boundary within and between words. After having demonstrated that both cues to segmentation were used and integrated by nine-month-olds Mattys and colleagues pitted the cues against each other. This served to test whether infants gave the same weight to both cues during word segmentation. In this follow-up study infants were found to listen longer to strong-weak stimuli that violated phonotactic constraints than to weak-strong words that did not. Mattys et al. concluded that nine-month-olds rely on prosodic cues as a first pass, which is then supplemented by other cues, including phonotactics.

Finally, Johnson and Jusczyk (2001) were the first to investigate the relative strength of statistical cues to word segmentation. Having replicated Saffran et al.'s (1996a, b) demonstration of infant sensitivity to statistical cues with natural speech, Johnson and Juczyk then explored eight-month-olds' relative use of statistics when the statistical cues were pitted against stress and coarticulation respectively. They found that both speech cues were weighted more heavily than statistical cues.

To conclude, infants have been shown to be sensitive to multiple cues to word segmentation. Moreover evidence suggests that infants begin to integrate some of these cues from approximately nine months of age. Cue integration is essential because reliance on a single cue is not sufficient for parsing the input correctly. More research is required regarding infants' ability to use multiple cues at different stages in development and how their weighting may change at critical points (e.g. during lexical acquisition).

1.2.5 Link between phonetic category formation and word recognition

It has been shown that infants become sensitized to the sound structure of their language during the course of the first year of life. They thus possess considerable language-specific knowledge before learning what words mean. However, studies of infants' phoneme-level representations have largely employed discrimination and generalization paradigms using isolated syllable tokens. (e.g. Grieser & Kuhl, 1989; Kuhl, Williams, Lacerda, Stevens & Lindblom, 1992; Werker & Tees, 1984). Consequently, the relevance of these results for vocabulary development has been disputed because such tasks may underestimate the difficulties of language processing under natural conditions.

Hence at least some controversy remains regarding the amount of detail present in the developing lexicon and the extent to which early representations are language-specific. This controversy is linked to the debate about whether there is continuity between infants' detailed and language-specific phoneme-level representations and their early word form representations.

Among the early work addressing this question were studies by Shvachkin (1973) and Barton (1976, 1978). Shvachkin (1973) taught young Russian children novel word forms for unfamiliar objects and subsequently tested children's discrimination of the newly taught words by asking his participants to hand him the associated objects. It was not until the end of the second year of life that Shvachkin's participants were able to perform this discrimination task, which led Shvachkin to conclude that younger children's early word form representations did not yet contain sufficient phonetic detail to allow them to distinguish the minimal pairs. Other researchers, including Barton (1976, 1978), reported findings suggesting considerable confusion of minimal pairs by two-year-olds, using a picture pointing task. Even between 27 and 35 months children were still only able to distinguish those words that they had used productively but not the ones that had been taught as part of the study

More recently, Stager and Werker (1997) also reported lack of early phonetic specificity in 14-month-old novel word learners using the 'Switch' design, originally designed by Cohen (1992) and subsequently modified by Werker, Cohen, Lloyd, Casasola & Stager, 1998). Following habituation to two word-object pairings, it was tested whether infants would detect a switch in the pairing (Werker, Cohen, Lloyed, Casasola & Stager, 1998). When 14-month-olds were required to make associative word-object links of the two similar-sounding word forms 'bih' and 'dih' they failed, even though infants were able to discriminate the words (Stager & Werker, 1997). As it was subsequently shown that 14-month-old children still fail to access the phonetic detail when well-formed English words are used (i.e. pin and bin) (Pater & Werker, in press), the abovementioned finding cannot be the result of the phonotactic violation of 'bih' and 'dih' not being possible words of English because they end in a lax vowel.

In contrast to the findings reported above, Juszcyk and Aslin (1995) and Swingley and Aslin (2000, 2002) presented evidence supporting the notion that infants do have access to phonetic detail in their early lexical representations from as early as 7.5 and 14 months respectively. Both sets of studies made use of a word-recognition test paradigm: Specifically, Jusczyk and Aslin familiarized infants with word forms in isolation (including 'feet' and 'bike') and subsequently presented the target words embedded in short text passages either containing the familiarized words or similar sounding foils. Differential listening time to familiar and unfamiliar texts was measured. Seven-and-a-half-month-olds were found to listen longer to familiar passages, thus demonstrating that they had recognized the words in the passages. Even when target words were replaced by minimally different nonwords, infants still demonstrated longer listening times to familiar stimuli. Given that targets and foils were minimal pairs, it was concluded that infants must have processed the target words in some detail for them to recognize them in the passages. Swingley and Aslin (2000) presented 18-23 month-olds with pictures of familiar objects on a TV-screen and presented either correct pronunciations or mispronunciations of one of the objects on the screen while monitoring infants' eye movements. Immediate effects of the mispronunciations were evident in that infants spent more time looking at the correctly named object and also required less time to fixate on the target picture during the 'correct pronunciation' condition. No effect of vocabulary size was observed. These findings were taken to suggest that 18-23-month-olds' early lexical representations are well-specified phonemically.

See also Swingley (2005b), for a related demonstration in Dutch eleven-month-olds using a different method; and Swingley (2003), for a demonstration of children's sensitivity to word-medial mispronunciations.

How can these seemingly contradictory results be reconciled? It has been suggested that a possible reason for the controversy regarding the level of phonetic detail in infants' early lexical representations may be methodological in nature (Stager & Werker, 1997; Swingley & Aslin, 2000; Swingley & Aslin, 2002; Fennell & Werker, 2004). Specifically, task demands differed considerably across the studies described above: Shvachkin's study required infants to learn new to map new words onto objects and make an explicit behavioural response during testing. Similarly, the experiment by Stager and Werker (1997) involved infants making novel word-object associations, so that it required considerable computational resources (Werker & Fennell, 2004).

In contrast, the studies by Swingley and Aslin (2000, 2002) and Swingley (2003) involved infants' eye-movements being monitored online, as opposed an explicit offline response being demanded from infants. Thus this task presumably demanded fewer computational resources even though it was still referential in nature. Task demands were also lower in the abovementioned study by Jusczyk and Aslin: here, infants were not required to access the meaning of the word forms (Fennell & Werker, 2004), nor were infants required to form word-object associations. Instead, infants had to recognize a familiar word form and then distinguish it from a phonetically similar word form.

Recently, Fennell and Werker (2003a,b) demonstrated that both explicit word knowledge as well as word-object familiarity allow 14-month-olds to access fine phonetic detail in the Switch task. Infants' success in this task was attributed at least in part to the lowered task demands under these conditions.

To conclude, infants appear to have access to phonemic detail under at least some task conditions (e.g. when applying their sensitivity online or when using a familiar word form), whereas their sensitivity may be masked when resource demands are higher (e.g. when an explicit response is required or when exclusively novel word forms are used).

Importantly, the studies discussed in this section involved primarily consonant manipulations, as have the majority of studies of this nature. However, it was pointed out in section 1.1.1 that language-specific attunement to vowels has been shown to precede that of consonants. Possible reasons behind this will be addressed in the following section as will a further challenge for infants as they learn to process vowels in language-specific ways: namely, paying attention to relevant and discarding irrelevant phonetic variation.

1.2.6 Learning about vowels

Compared with consonants, realizations of vowels involve less clearly defined articulatory gestures since there is no specific point or area of contact. Vowel realizations are thus uninterrupted by vocal tract constrictions and hence continuous segments. Perhaps therefore, vowel discrimination functions tend to be more gradual and their slope not as abrupt/steep as that of consonants. General parameters used to describe the realization of vowels are highness or openness (i.e. whether the lips are close together or far apart), roundedness (i.e. whether the lips are rounded or spread) and frontness (i.e. to what extent they are realized at the front of the oral cavity). Acoustic parameters employed to specify vowel identity are the first two formant frequencies F1 and F2 and for some vowels such as /i/, the third formant F3 also tends to be included.

In addition, vowels are always voiced in their canonical form and tend to be longer in duration than consonants (Polka & Werker, 1994). Elongation of vowels seems to be particularly common in speech directed to infants (e.g. Fernald, 1984). There is evidence suggesting that attunement to language-specific aspects of vowels precedes that of consonants (Kuhl et al., 1992; Polka & Werker, 1994).

Specifically, Kuhl et al. (1992) explored the influence of the ambient language on the internal structure of infants' vowel categories. The present study was based on Kuhl's earlier demonstration of the so-called Perceptual Magnet Effect (PAM) (Kuhl, 1991). According to Kuhl (1991), the Perceptual Magnet Effect arises when one listens to speech sounds of different prototypicality: the sound that is judged to be the best or most prototypical example of a particular sound category acts like a magnet in that it pulls less prototypical tokens closer to it perceptually than would be the case if perception were merely based on acoustic similarity space. A prototypical vowel is a vowel that is judged by native listeners as being representative or typical of that vowel in the native language in terms of how it sounds.

To examine the influence of the ambient language on the internal structure of infants' vowel categories at six months, Kuhl et al. (1992) created two sets of synthetic vowel tokens (i.e. 'peripheral vowels') around a 'central prototype'. The first set of vowel exemplars was synthesized around the English vowel /i/ and the second set around the Swedish vowel /y/. Both are high front vowels; /i/ is produced with spread lips and /y/ with rounded lips. The individual tokens within each set of peripheral /i/ and /y/ vowels varied from the central vowel token along equal mel steps of the first and second formant. The study revealed language-specific differences: six-month-old English infants showed an increased magnet effect (generalizations of peripheral vowel to central vowel) when the central vowel was the native English vowel as opposed to the non-native Swedish vowel, whereas the pattern was reversed for a group of Swedish infants. Kuhl and colleagues concluded that the native language affects the internal structure of at least some vowel categories from as early as six months and hence at an earlier point in development than for consonants.

Further evidence that attunement to language-specific aspects of vowels precedes that of consonants was presented by Polka and Werker (1994). Polka and Werker examined changes in discrimination performance of non-native vowel contrasts across the infancy period. The study consisted of a comparison of English infants' discrimination of two pairs of non-native German vowel contrasts (dYt versus /dut/ and /dyt/ versus /dut/). A native English contrast was also included as a control to ensure that infants could perform the task (/dat/ versus /dit/). Infants were tested at four and at six-to eight months, using the habituation-dishabituation looking procedure.

Findings suggested the presence of a developmental change in discrimination of non-native vowel contrasts between four and six months of age: Whereas four-month-olds discriminated both non-native vowel contrasts, six-month-olds discriminated neither of the two, thus providing evidence for a decline in discrimination of non-native vowel contrasts from a language-general to a language-specific pattern (Polka & Werker, 1994).

There is also evidence suggesting that six-month-old infants can successfully filter out relevant cues to vowel identity when stimuli vary along irrelevant dimensions (i.e. pitch and speaker identity) at the same time. In one study, infants of five and a half to six and a half months were presented with synthetic tokens of that kind using the conditioned headturn paradigm (Kuhl, 1979). The stimuli simulated an adult male, female and child talker. There were two stimulus tokens within each talker category, one with a fall-rise and one with a rising pitch. Infants learnt to ignore the pitch variation and hence successfully responded when the background stimulus /ɑ/ changed to the foreground or reference stimulus /i/ irrespective of its accompanying pitch changes and even though training had involved only a single exemplar of the two stimulus types and hence no pitch variation.

Interestingly, Kuhl and Hillenbrand (1979) subsequently demonstrated that infants' 'failure' to attend to the irrelevant dimensions in the preceding study could not be explained by their simply not being able to process this source of variability in sufficient detail: Using the same paradigm, Kuhl and Hillenbrand showed that six-month-olds could be trained to sort vowel sounds on the basis of pitch contour while ignoring accompanying vowel quality changes. In their later study, infants learnt to make a head-turn response to a vowel with a rising pitch contour and to refrain from turning their head when same vowel had a falling pitch. This response pattern was even generalized to new tokens that followed the same pitch pattern.

To sum up, it appears that six-month-old infants treat vowel quality as more important than accompanying pitch changes as long as testing is not preceded by training against this tendency. When specifically trained to attend to variation of pitch, however, as in the study by Kuhl and Hillenbrand (1979), infants can switch their focus of attention. The fact that infants are even capable of categorizing new tokens according to the learnt pitch modulation is important in that it demonstrates that their sophistication in processing fundamental frequency shifts surpasses simple discrimination. Apart from six-month-olds' preference for vowel quality over pitch modulation changes, Kuhl's studies show that infants' processing of specific dimensions or acoustic parameters at six months is not exclusively determined by what is phonologically relevant in the ambient language. Perhaps the exaggerated prosodic variation of IDS serves to train infants to regard pitch changes as an irrelevant source of variation.

Thus the evidence suggests that infants' vowel categories are being shaped by the ambient language from as early as six months of age and thus prior to the age at which language-specific effects have been demonstrated for consonants (Kuhl et al., 1992; Polka & Werker, 1994).

There are at least three possible explanations or contributing factors as to why infants may develop native language sensitivity to vowels early:

First, infants' listening time to vowels is likely to be increased compared with that of consonants because of some of the characteristics that set vowels apart from consonants: for instance, their more continuous realization and the fact that they are always voiced. In contrast, there are voiced and voiceless consonants (acoustically, the difference is a difference in voice onset time or VOT). The more continuous nature of vowels in particular may allow infants to process them in more category-specific (and hence also language-specific) detail. There are consonants, though, that may be argued to fall between consonants and vowels regarding their continuity, namely liquids, glides and approximants (e.g. /l/, /r/, /j/ and /w/). /w/ and /j/ are also referred to as semi-vowels. It is not clear whether these sounds 'behave' any differently from obstruents, for example with respect to the point in time at which infants start processing them in language-specific ways. It would be conceivable that attunement to these sounds may also precede that of plosives.

A second potential contributing factor regarding early attunement to vowels may be infants' relatively early language-specific processing at the suprasegmental (i.e. prosodic) level, which was discussed above. The difference in the time-course of suprasegmental versus segmental development led Cutler and Mehler (1993) and Polka and Werker (1994) to hypothesize that vowels, which carry a considerable amount of suprasegmental information, may be attended to more than consonants (at least in the early infancy period). The latter may hence be at least a partial explanation for the difference in time-course of the development of vowels and consonants. Although it would have been conceivable that language-specific tuning of vowels and consonants occurs at a similar point in time (namely by ten months) because vowels also carry segmental information, this was borne out by Polka and Werker's findings (Polka & Werker, 1994).

The third possible reason as to why language-specific vowel perception may precede attunement of consonants is the influence of infant-directed speech (IDS). It has been demonstrated that the acoustic space of vowels is stretched in IDS, thus making these segments more discriminable and more intelligible (Kuhl, Andruski, Chistovich, Chistovich, et al., 1997). Although this phonetic-level modification of speech addressed to infants is accompanied by other modifications at the prosodic, syntactic and semantic level (e.g. Grieser & Kuhl, 1988; Cross, 1977; Snow, 1994; for a review, see Liu, Kuhl & Tsao, 2003), recent findings suggest that it is the phonetic modification that is associated with advanced phonetic discrimination skills in infants both at six to eight and at ten to twelve months (Kuhl, Tsao & Liu, 2003). It may be, therefore, that infants' exposure to exaggerated phonetic units is a contributing factor to early language-specific tuning of vowels relative to consonants. Although Kuhl et al. (2003) tested discrimination of consonants rather than vowels, the perceptual advantage for vowels may still have exceeded that of consonants because it is the vowel space that is particularly stretched in IDS.

1.2.7 Cross- linguistic variation in the role of vowel length

The studies on language-specific processing of vowels described above all looked at fine-tuning to native language vowel quality (i.e. spectral features) rather than vowel duration. However, there are languages for which vowel length is as much of a determining feature of vowel identity as vowel quality. Among these are Japanese, Finnish and Dutch (see section 3.1, for a more detailed discussion of the differing role that vowel length plays across languages). Just as infants acquire knowledge about the language-specific quality features that mark a phonological contrast in their native language, they also have to learn whether vowel length can signal a phonemic contrast. This has to be learnt because its role varies among languages.

There is a major complicating factor, however: Children acquiring a language such as English which does not have a phonemic vowel length contrast cannot merely ignore vowel duration when listening to English speech, since this variation plays a crucial role in another aspect of phonological processing. In English, vowel duration largely determines the voicing status of postvocalic plosives; the length of vowels preceding voiced consonants is almost twice that of vowels preceding voiceless consonants (House, 1961). This effect is not specific to English, of course. Although there are differences of degree across languages (Chen, 1970; cited in Buder & Stoel-Gammon, 2002) relatively few languages show no interaction between vowel length and consonant voicing. Among the latter are Polish, Czech and Saudi Arabic (Keating, 1985; cited in Buder & Stoel-Gammon, 2002). Languages in which vowel length is contrastive tend only to have relatively small vowel length differences in these contexts and listeners do not seem to attend to this source of variation for determining the voicing status of the following plosive as much.

Learning about the specific role that vowel length plays in one's native language is made even more challenging by the fact that the length of vowels also varies as a function of linguistic context. A range of contextual factors affect vowel length systemically across languages: syntactic context, stress and speech tempo (see 3.1).

Given this considerable variation both within and across languages, it is likely to be a considerable challenge for infants to learn to process vowel duration in a language-specific way. One of the main goals of this dissertation is therefore to explore infants' sensitivity to this duration-based opposition as they learn more about the specific role vowel length plays in their native language. Two cases will be compared: the acquisition of Dutch, and the acquisition of Canadian English. The fact that there is a phonemic duration-based opposition in Dutch, but not in Canadian English, makes these two languages an interesting test case.

1.2.8 Dutch and North American English vowels compared

The studies discussed in this dissertation focus on infants learning 'standard Dutch' (henceforth referred to as Dutch) and 'General Canadian English' (henceforth referred to as Canadian English). The vowel inventory of the two languages is now briefly described. Of particular importance will be the role that vowel length plays in the two languages.

1.2.8.1 Standard Dutch

According to Booij (1995), Dutch has 15 vowels, of which 12 are monophthongs and at least three are diphthongs. The monophthongs are /ɪ/, /ɛ/, /ɔ/, /ʏ/, /ɑ/, /i/, /y/, /u/, /e/, /ø/, /a/ and the reduced vowel /ə/. There are three diphthongs that seem to be realized as such by all native speakers of Dutch. These are /ɛi/, /œy/ and /ɔu/. The long mid-vowels /e/, /ø/ and /o/, on the other hand, may or may not be realized as diphthongs depending upon the individual speaker. Among the vowels there are some that can be classified as short or lax versus long or tense (see Figure 1.1) (Booij, 1995, Gussenhoven, 1999). It is important to note that the Dutch vowels classified as long are not always longer than those classified as short. Specifically, the high long vowels /i/, /y/, /u/ only have the length of the other long vowels /e/ and /a/ (i.e. approximately 200 ms) if they precede /r/ (e.g. vier, vuur). Otherwise their length is comparable to that of the Dutch vowels classified as short (i.e. on average 100 ms) (Nooteboom, 1972; also see Booij, 1995).

The Dutch Vowels	
Short Dutch vowels	/ɪ/, /ɛ/, /ɔ/, /ʏ/, /ɑ/
Long Dutch vowels	/i/, /y/, /u/, /e/, /ø/, /a/

Figure 1.1: Short and Long Vowels of Dutch

Short or lax Dutch monophthongs and long or tense Dutch monophthongs are subject to different phonotactic constraints: The five lax vowels /ɪ/, /ɛ/, /ɔ/, /ʏ/, /ɑ/ are more restricted phonotactically in that they do not typically occur as part of an open syllable (i.e. they are always followed by at least one consonant, as in /tɑk/). The sound sequence /tɑ/, therefore, is not allowed in Dutch. The long Dutch vowels /i/, /y/, /u/, /e/, /ø/, /a/ and the diphthongs can occur in any syllable type (i.e. open or closed). Therefore, the sound sequence /ku/, for example, is permissible in Dutch (it also happens to be an existing Dutch word).

There are several Dutch vowels that have been referred to as 'marginal' because they occur only in loanwords. Among them are /ɛ:/ (e.g. enquete), /ɔ:/ (e.g. zone) and /œ:/ (oevre) (Booij, 1995). Gussenhoven classified a number of additional vowels as marginal, namely /i:/, /y:/ and /u:/ (Gussenhoven, 1999).

1.2.8.2 Canadian English

'General Canadian English' refers to the dialect spoken in western and central Canada. This Canadian dialect is described as having eleven monophthongs, of which five are lax vowels and five tense vowels (Avis, 1975; Gregg, 1975; Walker, 1975; see also Morrison, 2000). The eleventh vowel is the reduced vowel schwa /ə/; its quality varies considerably in different contexts (Ladefoged, 1999). The lax vowels are /ɪ, ɛ, æ, ʌ, ʊ/ and the tense vowels /i, e, ɑ, o, u/. Although it has been suggested that there is not always a difference in duration between tense and lax vowels (specifically between /i/ and /ɪ/) in the type of Canadian English spoken in Vancouver (Gregg, 1975), this does not appear to apply to all speakers nor do there appear to be experimental data to corroborate this claim (Morrison, 2002).

The Canadian Vowels	
Lax Canadian vowels	/ɪ, ɛ, æ, ʌ, ʊ/
Tense Canadian vowels	/i, e, ɑ, o, u/

Figure 1.2: Lax and Tense Vowels of Canadian English

Despite the likely presence of accompanying durational differences between tense and lax vowels in at least some Canadian speakers' vowel realizations, spectral cues appear to serve as the primary cue to vowel identification for Canadian English listeners (e.g. Assman, Nearey & Hogan, 1982; also see section 3.1).

There are also at least three diphthongs in Canadian English: /aɪ/, /ɔɪ/ and /aɪ/.

It should be noted that the above classification is oversimplified in the sense that a number of shifts have been observed in different variants of Canadian English, including vowel realizations of Vancouver speakers. Hence this classification merely serves to provide a general overview of Canadian vowels.

1.2.9 The challenge of learning about the role of vowel length in the ambient language

Based on the differing role of vowel length in Dutch and Canadian English it follows that Dutch-learning infants must learn to pay attention to vowel length for the purpose of vowel identification for at least some vowels of Dutch, whereas Canadian-learning infants should eventually learn to ignore vowel length, at least for the purpose of determining vowel identity.

This dissertation aimed to explore how and when Dutch learning infants learn to pay attention to vowel length cues and when English learning infants learn to pay relatively less attention to vowel length variation for making lexical distinctions.

It is possible that vowel length may be more salient acoustically than some non-native vowel quality differences. It is known from other research that infant listeners are very sensitive to temporal cues; for instance, infants, like adults, are able to correctly categorize phonemes despite rate of speech differences (Miller & Eimas, 1994). It is thus conceivable that young listeners are similarly sensitive to temporal cues in other aspects of psycholinguistic processing, such as when categorizing phonemes during the initial stages of word-learning, even if the temporal contrast is not meaningful in their native language. As a result, such listeners may at least still be able to tell the two sounds apart if presented with both the long and the short vowel sound in linguistic context. If this were correct, it would be expected that infants whose native language does not have a long-short vowel opposition may perceive two word forms differing only in the length of the vocalic segments as two different words rather than mapping them onto the same word form. Alternatively, listeners may be able to distinguish the two word forms without going as far so as to categorize them as distinct lexical items in the sense that they may interpret the word containing the lengthened vowel as a less prototypical realization of its short vowel counterpart.

One study that has explored the processing of vowel length in adult subjects suggested that at least those listeners whose native language has a phonemic vowel duration contrast make use of their duration-based listening strategy when listening to a second language (henceforth L2). Japanese learners of English applied – or misapplied – their sensitivity to vowel duration when processing non-native vowels of English such as /i/ and /ɪ / for which vowel duration is not only not the most salient but also not the most reliable cue to identifying vowel identity (Morrison, 2002). Japanese adult listeners, who were all highly proficient in English, were asked to categorize CVC word forms containing a range of vowels, including /ɪ/ and /i/. Confusion matrices revealed that Japanese subjects' rate of identification of English /ɪ/ was almost as high as those of native English speakers (averaged over voiced and voiceless CVC-contexts). However, Japanese listeners' identification of English /i/ was much lower overall. Interestingly, the effect was carried entirely by the much decreased percentage correct where /i/ was followed by a voiceless consonant, presumably because Japanese speakers had misapplied duration cues to the extent that they had categorized the much shorter duration of the English /i/ in the voiceless context as the short English /ɪ/. When Morrison measured the durations of the vowels in the different contexts and compared them with the average durations of the Japanese native vowel /i/ (short) and /iː/ (long) and with Japanese listeners' duration thresholds for these two vowels in CVC isolated word contexts, he corroborated his hypothesis that Japanese listeners had primarily used their native long-short vowel duration threshold when identifying English /ɪ/ and /i/ (Morrison, 2002). The much decreased identification of English /i/ in the voiceless context and the much higher identification of English /ɪ/ in both voicing contexts could also be explained in that way since only English /i/ in the voiced context reliably fell within the range of the Japanese long /i:/. Some use of spectral information was apparent both in Morrison's study and in other comparable published studies (e.g. Strange, Akahane-Yamada, Kubo, Trent, Nishi & Jenkins, 1998; Ingram & Park, 1997).

The lack of reliability of the use of that cue notwithstanding, it may still serve a purpose for Japanese listeners to use vowel duration cues as opposed to vowel quality cues in that Japanese listeners have a lot of practice with the use of that cue – because vowel duration is contrastive in Japanese. Moreover the use of the vowel duration cue clearly helped them to some extent when identifying the non-native English sounds.

However, as the above study exemplified, there are conditions under which drawing upon the Japanese temporal-based opposition is not sufficient. For example, when the target vowel is followed by an obstruent, vowel length tends to vary as a direct function of the voicing status of that obstruent both for lax and for tense vowels. That variability then tends to occlude the reliability of the duration cue, in that it may increase the duration of /ɪ/ such that its length overlaps with /i/ in other contexts. Therefore, the Japanese listening strategy was less than ideal when categorizing English vowels.

It should be noted that similar cases of using native listening strategies in non-native listening have been observed by Weber (2001) for phonotactics and by Broersma in lexical decision and categorization (Broersma, 2005a). In a recent experiment by Broersma (2005b), it was shown that English listeners' voicing judgements of the syllable-final plosives (/v/ and /f/) were affected when the length of the preceding vowel was ambiguous. Specifically, English listeners had difficulty judging the syllable-final fricatives as voiced when the preceding context had been extracted from a word ending in the voiceless variant of the two consonants and vice versa. Interestingly, Dutch listeners' voicing judgements were unaffected by the informativeness of the preceding vowel. The latter resulted in the Dutch outperforming the English in their voicing judgements in the ambiguous conditions, even though there is no voicing distinction of syllable-final fricatives in Dutch. Presumably the fact that Dutch listeners' native phonology had not biased them to pay attention to vowel length in this particular context allowed them to attend to the very small acoustic differences between the voiced and the voiceless variants of the two fricatives.

This introduction served to elucidate challenges that infants face as they learn the most relevant rhythmic, distributional and phonological properties of the ambient language, while at the same time having to learn to ignore a substantial amount of irrelevant acoustic variation in the signal.

In this research project the goal was to explore one source of variability, namely vowel length. As we have shown, vowel length provides a particularly interesting test case in that it is phonemic in some but not all languages, while it can still play an important role even when it is not phonemic. By comparing infants' processing of vowel length cross-linguistically in two groups of infants for whom vowel length is either phonemic or non-phonemic, we can observe how infants' perceptual sensitivities develop into language-specific phonological knowledge.

2 DUTCH INFANTS' SIMILARITY JUDGEMENTS OF VOCALICALLY VARYING WORD FORMS IN UTTERANCES

2.1 Introduction

In studies employing discrimination and generalization paradigms using isolated syllable tokens, infants have been shown to be sensitive to language-specific detail towards the end of the first year and prior to the onset of word-learning (e.g. Werker and Tees, 1984; Kuhl et al., 1992; Polka & Werker, 1994). While it is undoubtedly valuable to look at young infants' perception of speech sounds and syllables in isolation, the relevance of such findings to word learning remains questionable. Hence, controversy remains over the degree of phonetic detail present in the developing lexicon (Jusczyk, 1997; Hallé & Boysson-Bardies, 1996) and the extent to which early representations may be language-specific, although research regarding the first question has been emerging, as discussed in section 1.2.5 (e.g. Stager & Werker, 1997; Swingley, 2001; Swingley & Aslin, 2002; Fennell & Werker, 2004). This dissertation addresses both these questions.

Infants of seven and a half months are able to recognize words in sentences both in preference paradigms (Jusczyk & Aslin, 1995) and in conditioned headturn tasks (Morgan, personal communication, 2001). More recently, electrophysiological evidence for ten-month-old infants' incremental recognition of familiar word forms in continuous speech has also become available (Kooijman et al., 2005). Specifically, Kooijman and colleagues familiarized a group of Dutch infants with lists of different tokens of novel words in isolation. These were low-frequency Dutch words unlikely to have been known to infants. Subsequently both the familiar (as confirmed by measuring ERP responses during familiarization) and novel words were embedded in sentences to ascertain whether infants would recognize the embedded word forms and to obtain detailed information about the time course of infants' early word recognition. Not only did infants recognize the embedded words, but infants' ERP responses also indicated that their recognition of the embedded words during testing tended to occur after as little as half a second after the words' onset, before the word ended. Word recognition was thus incremental, just as in adults (Frauenfelder & Floccia, 1998) and in toddlers (Fernald, Swingley & Pinto, 2001; Swingley, Pinto & Fernald, 1999).

To conclude, these ERP results suggest that ten-month-old infants are able to recognize familiarized word forms in continuous speech when they have been presented with merely half a second of the signal. In light of the present research questions, it would be interesting to extend this line of research by including phonetically similar word forms to ascertain at what point in the signal infants are able to disambiguate the phonetic foil from the target word.

Overall, it appears that infants are indeed capable of recognising trained word forms (and even parts thereof) in sentences before they start producing their first words. Moreover, Jusczyk and Aslin (1995), using a familiarization procedure, provided evidence consistent with accurate phonological specification in such words: infants did not tend to generalize from the training word to a minimally different control word. For example, children trained to make a headturn response to the word 'cup' did not display the trained behaviour to the minimally different word form 'tup' (Jusczyk & Aslin, 1995). The word form 'tup' differs from 'cup' in one phonemic feature only, namely its place of articulation (i.e. velar versus alveolar). Infants' ability to categorize these two word forms as distinct suggests that they store word forms taught in the laboratory with considerable phonetic detail. However, the study does not speak to whether or not the word forms that infants acquire in the laboratory are being fit into a perceptual system that is tuned by the native language phonology: Jusczyk and Aslin's control words all differed phonemically from the training word. To our knowledge, there are no studies that have looked at language-specificity of infants' representations of early word forms, let alone of minimally differing words imbedded in utterances. Such a study would address the question of whether or not infants are capable of drawing upon language-specific phonemic representations when acquiring their first word forms.

To determine to what extent low-level phonetic or acoustic variation is used during the processing of continuous speech Experiment 1 explored Dutch infants' implicit similarity judgments to word forms embedded in utterances, while the word forms differed vocalically, either phonemically or non-phonemically. It is conceivable that the attentional resources required for the processing of utterances, as opposed to isolated word forms, will interfere with infants' fine-grained processing of the vowel segments. At the same time, it may be that infants will start using their special sensitivity to native language vowel information when first recognizing words in fluent speech.

2.2 Infants' similarity judgements of word forms in utterances

Experiment 1 was designed to explore infants' judgements of the similarity of vocalically varying word forms in utterances. We were interested in infants' sensitivity to native and non-native variation of a trained target word embedded in short sentences. It should be noted that the experiment did not look at infants' use of language-specific phonemic representations when acquiring word-object associations, since the conditioned headturn task does not require infants to map the word forms onto pictures or objects.

Infants were first trained to make a head-turn response to the word form 'hasp' and subsequently presented with utterances containing either the trained target 'hasp' (e.g. 'Jullie hebben een fijn hasp') or one of three variant word forms containing a different vowel (e.g. 'Jullie hebben een fijn haasp'). We were interested in infants' headturn responses to the training word relative to its variants.

Infants ranged in age from seven and a half to twelve months, so we could be fairly confident that our participants would be able to perform the task of spotting the trained word form in the sentences.

2.3 Method

2.3.1 Participants

Subjects were sixteen Dutch infants ranging from seven and a half to twelve months. 16 additional infants were tested but not included in the final sample because of crying, fussiness, failure to meet the training criterion or fear of the reinforcer. Infants were drawn from city birth records. All infants were without apparent health problems. Exclusion criteria included hearing loss, a history of ear infections and multilingual home environments. All participants were monolingual Dutch infants exposed to Dutch at least 90% of the time (by parental report). All participating infants' parents received a small payment to thank them for their efforts.

2.3.2 Materials

The stimuli consisted of ten isolated words, including the target, and a range of short, simple utterances. The materials are listed in full in the Appendices in section 7.1. The stimuli were spoken by a bilingual female native speaker of Dutch and American English. Four of the isolated words that were used as background stimuli (see Procedure on page 25) in the training phase matched the target word 'hasp' with respect to its consonantal context either word initially (e.g. 'hets') or word-finally (e.g. 'tesp'). This was intended to increase infants' attention to the vocalic portion of the word. The target word was always sentence-final because previous research indicated that seven-and-a-half-month- olds cannot reliably detect words in sentence-medial position (Morgan, personal communication, 2001). Moreover there was a phonotactic word boundary cue between the last word of the carrier sentence and the target word to facilitate segmentation: the carrier sentences always ended in the sound sequence –ijn. The following target word or its variant began with either /f/ (control word) or /h/ (all other words).

To ensure that infants would not be distracted by words they already knew in the carrier sentence, words commonly known to infants at this age were avoided. The target vowel /ɑ/ was not present in the carrier phrases. To prevent carryover effects, each carrier sentence was only used once with each target word.

2.3.3 Procedure

A variant of the Conditioned Headturn Paradigm (henceforth CHT) was employed for this experiment. The experiment was controlled using NESU (Nijmegen Experiment Set-Up) experiment control software, developed at the Max Planck Institute for Psycholinguistics (MPI) in Nijmegen.

The Conditioned Headturn Paradigm is a behavioural method that - as used in this project - involved infants being trained to respond to the change of a range of word forms such as 'tesp' and 'pesk' to a specific target word, namely 'hasp'. As may be clear from the above example, infants were not required to perform a discrimination task but rather a categorisation task. More commonly, the paradigm is used in a different way in that infants are taught to make a head-turn response to a visual reinforcer when a change occurs from one speech sound to another. Hence, it is traditionally considered a discrimination task and not a categorisation task (see Kuhl, 1985; Werker & Tees, 1984).

Our rationale for selecting the CHT-procedure was as follows: The primary advantage of employing the CHT-paradigm is that it yields richer individual-participant data than other procedures available for testing infants, such as the Headturn Preference paradigm or the Habituation procedures. For example, one can consider the number of trials infants require to acquire the conditioned head-turn response when looking at individual differences. Alternatively, first trial data can be compared with total-trial data. By comparison, the High Amplitude Sucking Paradigm is more limited in that respect because group data is required for conclusions to be drawn about infants' discrimination skills. The same applies to the Head-Turn Preference Procedure. In the latter subjects' listening times to two different types of stimuli are compared. Significant differences between conditions and hence infants' preference for one set of stimuli are interpreted in terms of subjects' differential sensitivity to certain aspects of the stimulus set in that condition (for more details see Kemler Nelson, Jusczyk, Mandel, Myers, et al., 1995).

Moreover, in contrast to habituation methods, the stimulus and the reinforcer are independent in the CHT-paradigm (Kuhl, 1985). Infants learn to produce a head-turn in response to a novel auditory stimulus, which is then reinforced visually with a stimulus known to be interesting to infants at this age. In contrast, sounds that are not reinforcing to infants in the High Amplitude Sucking-procedure could, in principle, yield no sucking response and thus a null result, even though infants may be able to discriminate them from another set of stimuli (Kuhl, 1985). Similar concerns apply to the headturn preference procedure.

Additional merits are that the procedure is versatile. It has been used to test speech perception of native and non-native contrasts in infants but also in subjects across the life span, albeit with minor modifications for non-infant participants (for details, see Werker, Shi, Desjardins et al., 1998). The traditional procedure is considered most useful for infants between five and a half and 18 months (Kuhl, 1985; Werker et al., 1998).

There is also evidence that the CHT-paradigm may serve as a tool for identifying children at risk for hearing loss or perceptual difficulties (Werker et al., 1998). A related finding is that the CHT-paradigm may have potential for assessing the relation between infant speech perception at six months and their early communicative development. Liu, Tsao, Stevens and Kuhl (1999) found significant negative correlations between the number of trials six-months-olds required to meet the criterion in the conditioning phase and 13-months-olds' comprehension of phrases and words. Moreover, Tsao, Liu and Kuhl (2004) used the CHT-method to demonstrate a strong pattern of correlation between early speech perception at six months (i.e. vowel discrimination) and later language abilities (i.e. word comprehension and production and comprehension of phrases). Finally, Kuhl, Conboy, Padden, Nelson and Pruitt (2005) employed the CHT-method in a study which revealed that a seven-month-olds infant's native language phonetic perception skills are strongly predictive of that child's later language development in terms of speed of language growth. Interestingly, the opposite pattern was found for early non-native speech perception skills: the better an infant's early non-native speech perception skills, the slower is that child's later language advancement going to be. It seems, therefore that Conditioned Headturn paradigm has considerable potential to explore the link between early speech perception and later language development (Kuhl et al., 2005).

2.3.3.1 CHT-methodology

The experiment took place in a quiet room. Infants were seated at a small table on their parent's lap across from the experimenter. Only one experimenter was involved in our variant of the procedure. The experimenter's task was to distract infants with a range of silent toys to maintain their attention and to prevent them from being directed towards the reinforcer throughout the study. At the same time, the experimenter operated two foot-pedals whose role will be described below. Sessions were video-recorded to allow additional off-line coding. Both the experimenter and the infant's parent wore headphones with masking music at all times to prevent the adults from having any knowledge of the stimuli that might lead them to bias the infants' behaviour. Speech stimuli had been superimposed onto the music so that there would be no silences during which the stimuli could be identified.

There were two phases, a training phase/conditioning phase and a test phase.

Training/Conditioning phase

The training phase began when infants were oriented towards the toys held by the experimenter. During this phase, infants were presented with several 'background stimuli' such as the word form 'tesp' at an ISI of 1200ms. An experimental trial consisted of two presentations of a 'novel stimulus' (i.e. the target word form 'hasp'). The target stimulus was initially paired with the 'visual reinforcer', which was activated simultaneously (i.e. at onset of the target word). As in previous implementations of the conditioned headturn procedure, the visual reinforcer (VR) consisted of two animated toy animals inside a dark custom-made Plexiglas box so that the animals would not attract attention unless the reinforcer had been turned on. On activation, the box was illuminated and one of the toy animals displayed (i.e. either 'Tigger' from the 'Winnie the Pooh story' or a small white dog). As the reinforcer toy moved and made noise appealing to infants, they spontaneously turned towards it. However, as one of the reinforcers ('Tigger') turned out to be frightening for a small number of the younger babies we activated only the light without accompanying sounds and movement for a few trials at the beginning of the study to allow infants to get used to and enjoy the reinforcement. Following Kuhl (1985), the visual reinforcer was accompanied by verbal praise (Kuhl, 1985).

Gradually, a delay between the presentation of the target stimulus and the VR was introduced. This delay was increased across trials: after two trials the reinforcer was activated 400 ms later and after a further two trials the reinforcer was activated another 400 ms later until a delay of 1600 ms had been reached. Eventually, infants learned to anticipate the reinforcer by turning their head to the target stimulus (i.e. the loudspeaker that emitted the stimulus) before the reinforcer had been activated (anticipatory headturn). Once subjects had reached our predefined criterion of five out of six consecutive correct headturn responses during the training phase, the test phase began.

Figure 2.1 summarises what happened during the training phase: infants were taught to make a headturn response to different tokens of the possible Dutch word form 'hasp', which contained the Dutch vowel /ɑ/ when presented among other 'background words'. By presenting a range of phonemically diverse background words we tried to ensure that infants would pay attention to the specific word form of the target word rather than merely learning to turn their head to any sound change. By sometimes using words with the same onset or offset as the target word we attempted to ensure that infants would pay particular attention to the vowel of the target word rather than just its onset or coda.

Infants were trained to turn their head when they heard the word form hasp, containing the Dutch vowel /ɑ/

e.g. pesk tesp hierk /hɑsp/ /hɑsp/

⬇ ⬇ ⬇ ⬇ ⬇

background words target word

headturn was trained here

Figure 2.1: Training phase

Discrimination/Test phase

During the test phase infants heard sentences containing one of the following: the trained target word form 'hasp', similar sounding variants of the trained target and dissimilar sounding controls. The test and control trials were quasirandomly ordered according to a new pre-determined trial order for each infant. A randomization programme was employed for making the randomizations. On test trials, the embedded word was the target word 'hasp' and infants' correct headturn responses to this word were reinforced. On control trials, the embedded word was one of three possible variant word forms of the target word and infants' headturn behaviour was monitored but not reinforced. The variant word forms were /ha:sp/, /hʌsp/ (containing a non-native vowel) and /fi:m/ (see Table 2.2). The controls served to assess the baseline probability of infants turning their head. The experimenter's task was to activate a left foot-pedal as soon as a headturn occurred. The main difference between the training phase and the test phase was that reinforcement was contingent upon the child having made a correct headturn response during testing but not during training.

We were primarily interested in infants' spontaneous generalization of the trained word form when they were presented with vocalically varying word forms embedded in utterances during the test phase (see Table 2.2 for an example sentence). Generalization was tested using the long native vowel /a:/ and the short non-native vowel /ʌ/. Infants were thus presented with utterances containing the trained /hɑsp/ or one of its variants (i.e. either /ha:sp/ or /hʌsp/). There was also a control condition during which 'fiem' was embedded in utterance-final position. The control word was to serve as a baseline to assess infants' tendency to turn their head to a word that has no phonemic resemblance with the trained word form.

Example sentence during test phase of CHT-study	
Onze beste leraar is Heijn /hɑsp/.	trained target
Onze beste leraar is Heijn /ha:sp/.	native vowel
Onze beste leraar is Heijn /hʌsp/.	non-native vowel
Onze beste leraar is Heijn /fi:m/.	control

Table 2.2: Test phase

We measured both traditionally used proportions of headturns across conditions and response latency.

2.3.3.2 Possible outcomes

On test trials hits and misses were assessed. A 'hit' refers to a correct headturn response to a change stimulus. A 'miss', in contrast, is a failure to turn during a change trial. An incorrect headturn during a control trial is referred to as a 'false positive' and correctly refraining from displaying a head-turn during controls is a 'correct rejection'. Only hits were reinforced.

The performance criteria by which infants are judged to have succeeded differ somewhat across studies. A criterion suggested by Kuhl (1985) is nine out of ten consecutive correct trials in the conditioning phase. It seems useful to include an additional criterion, namely a fixed number of trials, such as 30, within which the criterion must have been reached for training to be regarded as successful. Following Werker and Lalonde (1988; in Werker et al., 1998), we incorporated this additional constraint.

2.4 Predictions

We predicted that Dutch infants would respond most to the trained target word /hɑsp/ given that this is the word most directly associated with the reinforcer. Secondly, we expected infants to respond equally frequently to the word form containing the non-native vowel (i.e. /hʌsp/). The high response rate to the word forms containing the non-native vowel was predicted because of the phonological structure of Dutch and based on a pretest with Dutch adults demonstrating their conflation of /ɑ/ and /ʌ/. Third, we hypothesized that infants would respond less to the phonetically similar but phonemically distinct word form /ha:sp/. This prediction was based on the assumption that infants would perceive the trained target /hɑsp/ and the word form /ha:sp/ as two distinct word forms, based on the phonological structure of Dutch. Finally, we expected fewest responses to the control item /fi:m/ since it differs so radically from the trained target word that there should be no association with the reinforcer.

2.5 Results

The primary question to be addressed was whether infants would generalize their headturn response more readily to the non-native word form than to the native variant. Figure 2.3 shows proportions of headturn responses across conditions. Reaction times across conditions are displayed in Figure 2.4. On both the response proportions and response latency measures, infants more readily generalized from the trained target to the non-native word form than to the native variant. The effect was significant for response latencies showed a trend in the predicted direction for proportions.

A One-Way ANOVA yielded an overall main effect of word-type on the speed of infants' headturn responses ($F(3, 15)= 18.59$, $p<0.01$). Response proportions, however, did not differ significantly from one another ($F(3, 15)= .94$, $p= .43$, n. s.), although the trend was in line with our predictions, as is discussed further below.

Furthermore, planned comparisons for reaction times revealed that infants were significantly faster to respond to the trained target word 'hasp' than to the native variant 'haasp' ($t(15)= 2.66$, $p< .01$); all tests two-tailed). This was also the case for the control word 'fiem' ($t(15)= 21.56$, $p< .01$). Crucially, there was no significant difference in reaction time between the trained target word and the non-native variant word 'husp' ($t(15)= 1.04$, $p< .30$, n. s.), as predicted.

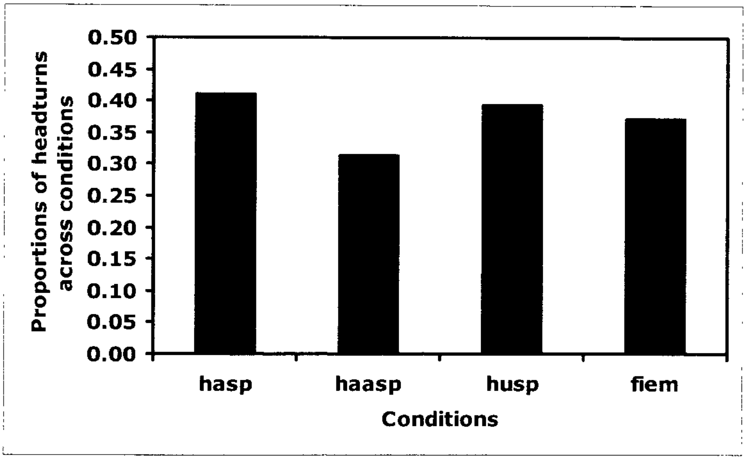

Figure 2.3: Headturn responses across conditions (Dutch CHT-experiment)

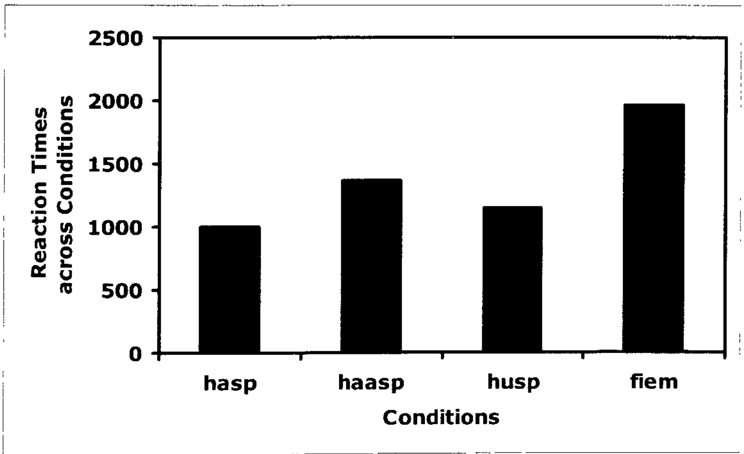

Figure 2.4: Response latencies across conditions (Dutch CHT-experiment)

2.6 Discussion

Experiment 1 revealed that Dutch infants' generalization pattern was in line with the phonology of the Dutch language. To recapitulate, infants more readily considered a word form as novel – in relation to the trained target word form /hɑsp/ – when it was the native variant /ha:sp/ as opposed to the non-native word form /hʌsp/. The effect was significant across the four conditions for response latencies though not for proportions. It should be noted that there is at least a trend for the proportions across the target, its native and non-native variant across conditions to be in line with our predictions: infants responded most to the trained target, closely followed by the non-native variant and finally by the native target.

Taking the proportion data in isolation, it is puzzling that there were more headturn responses to what was supposed to be the control word, namely 'fiem'; responses to 'fiem' exceeded those to the native variant /hasp/, although the latter is far more similar acoustically to the target than the control word. One possible explanation for this would be the presence of a 'novelty effect': 'fiem' is the only word form that is likely to 'pop out' acoustically and hence it may have attracted infants' attention for that reason. There is some support for the hypothesis that the control word was processed differently from both the target and its two variants, in that reaction time increased by almost 800ms when 'fiem' was presented. Infants may not have turned their head to it as fast as to the other words because they had learnt that it would not be followed by the reinforcer. They may still have turned to 'fiem' eventually, however, because it stood out acoustically. The fact that response latency to /ha:sp/ was much lower than to 'fiem', though higher than to the other variant words, indicates that /ha:sp/ was processed differently.

To conclude, our Dutch infants appeared to be less likely to categorize a word form as novel when it differed in a non-native way from an already known word (i.e. the training word) than when the difference was phonemic in nature, presumably because they assimilated the non-phonemic vowel /ʌ/ to the trained target vowel /ɑ/. Learning to ignore phonologically irrelevant variation is clearly beneficial for infants during word learning and beyond since it could help them to process the considerable natural variation of connected speech in a more efficient way by filtering out irrelevant variation (described in more detail in section 1.2.1).

Infants' pattern of generalization when responding to the native and non-native variant of the training word is consistent with Dutch phonology and with the hypothesis that infants between eight and 12 months are starting to become sensitive to language-specific aspects of vowels during word form recognition. However, there are at least two confounds that deserve further attention:

First, an acoustic explanation is also possible. This is to say, irrespective of Dutch phonology, /ɑ/ and /a:/ may simply be more discriminable or distinct than /ɑ/ and /ʌ/ since there is a considerable phonemic length difference accompanying the vowel quality difference in the former but not in the latter pair. To exclude an acoustic explanation of the current results, one would ideally test these contrasts on a group of infants for whom the phonology makes a different prediction. English-learning infants represent one such language group. If it were indeed native language phonology, which shapes infants' implicit word recognition judgements, a reversal of response latencies – in line with standard English phonologies – would be expected.

Second, while the procedure is useful for looking at discrimination and categorization, it only does so at a low level of processing. More specifically, the task might only assess infants' abstract similarity judgements. This may not allow conclusions to be drawn about a higher level of processing and hence the level of word-object associations.

2.6.1 Vowel length versus vowel quality

The previous experiment also raises a more general question, namely whether vowel length tends to be more salient acoustically than vowel quality differences. Why may this be the case? Vowel length is a temporal rather than a spectral cue and it could be easier for listeners to notice – even in cases where the native language does not contain a phonemic vowel length contrast. Supporting evidence for this notion was presented by Kondaurova and Francis (2004) who demonstrated that Russian listeners relied predominantly on duration cues when distinguishing the non-native English vowels /ɪ/-/i/ even though Russian has no phonemic vowel length contrast. The authors interpreted this finding as lending support to Bohn's 'desensitization hypothesis' (Bohn, 1995), which states that listeners distinguishing a non-native vowel contrast will employ vowel duration cues by default when spectral differences are insufficient. If this hypothesis it correct, the question arises whether this may be the case because vowel duration is perceptually easier to attend to even when its variation is non-phonemic for the listener.

An alternative interpretation would be that listeners for whom attention to vowel duration cues is found may have had to pay attention to vowel length variation for other aspects of speech processing in their native language and hence their sensitivity to this cue may reflect their experience with subphonemic duration contrasts within their native language. There is considerable evidence corroborating the notion that adult listeners are sensitive to and make use of subphonemic vowel duration cues during online spoken word recognition (e.g. Salverda, Dahan & McQueen, 2003; Crosswhite & Tanenhaus, 2004).

To our knowledge, no cross-linguistic studies on infants' processing of variation of vowel length during word learning have been conducted to date. Experiment 2 and 3 constitute a first attempt to address this question.

3 VOWEL DURATION IN A WORD CATEGORIZATION TASK

3.1 Introduction

In Experiments 2 and 3, we continued to address the question of which type of phonetic variation would signal a lexical contrast for infants and which variation would not, by exploring the nature of infants' generalizations in word form recognition.

Specifically, we focused on infants' processing of vowel length, comparing two subject populations, one for whom vowel length is phonemic and one for whom it is not. A native vowel quality contrast served as a control. As Dutch is a language with phonemic vowel length, whereas Canadian English is not, infants from these two populations were tested (see section 1 2.8 and discussion of the differing role of vowel length in English, Dutch and Japanese later in this section).

The method employed for both studies was the Conditioned Headturn Paradigm (CHT) (see section 2.3.3.1 for methodological details), just as in Experiment 1. However, there was a major difference between the two experiments: all stimuli were naturally produced in Experiment 1 (allowing for some quality differences between the vowels /ɑ/ and /a:/ in 'hasp' and 'haasp' respectively). In Experiment 2, however, length alone was manipulated to ensure comparable manipulations in Dutch and English. Experiment 2 was conducted at the Max Planck Institute for Psycholinguistics in the Netherlands with Dutch infants and Experiment 3 at the University of British Columbia in Vancouver, Canada with Canadian English infants.

Referring back to the processing of phonetic variation, learners ultimately have to discover that it is the segmental phonology of their language that governs which words contrast. The following studies explored whether 8-12-month old infants have learned whether vowel length can signal a phonemic contrast in their native language. This has to be learned because its role varies across languages: for example, in Japanese, double and single vowels contrast and hence the length of a vowel is critical for word meaning (i.e. it is phonemic). For example, two Japanese words that differ exclusively in this way are 'ojisan' (meaning 'uncle' and containing the single Japanese [i]) and 'ojiisan' (meaning 'grandfather' and containing the double Japanese [ii]). There are five double and five single vowels in Japanese with identical spectral properties (Akamatsu, 1997).

In English the length of a vowel does not determine vowel identity and hence vowel length is not phonemic. However, tense vowels of English nonetheless tend to be longer than lax vowels and hence English vowels do seem to have an intrinsic duration (House, 1961; Peterson & Lehiste, 1961). Moreover, as will be discussed further later in this section, vowel length is at least a contributing factor in determining the voicing status of postvocalic plosives and fricatives. It may seem, therefore, that vowel length is of similar importance as vowel quality in English. However, there is experimental evidence suggesting otherwise, although there appear to be at least some conditions that induce English listeners to increase their attention to vowel duration cues. Importantly, even in those cases vowel length cues do not serve to determine vowel identity. Three studies that have looked at this issue experimentally include Assmann, Nearey and Hogan (1982), Strange (1989) and Hillenbrand, Clark and Houde (2000).

First, the study by Assmann and colleagues indicated that American English adult listeners' vowel categorization judgements are not significantly impaired when listeners are presented with vocalic stimuli from which duration cues have been removed. This suggests that vowel length is not an essential cue for vowel categorization for American English listeners since categorization was still largely successful without the presence of the vowel length cues.

The study by Strange (1989) yielded more mixed results, however: whereas shortening 'silent-centre stimuli' such that the centre portion matched that of the shortest vowel nucleus did not affect listeners' identification judgements, lengthening them such that they matched the vowel nucleus of the longest vowel segment disrupted listeners' judgements. It is possible, however, that subjects' marked increase in error rate was the result of subjects perceiving the initial and final portion of the syllable as two different utterances rather than a single syllable. (Strange, 1989; cited in Hillenbrand et al., 2000).

Finally, in the more recent study by Hillenbrand and colleagues this interpretive problem was avoided by employing stimuli that had been modeled on speech. American English adult subjects were asked to identify synthesised CVC syllables generated at different durations. Stimuli were either synthesized using a formant synthesis technique or sinusoidally created. Both shortening and lengthening of syllables had a relatively small effect on intelligibility of the original vowels in that most signals were identified correctly both at their original durations and at all three altered durations. Interestingly, subjects' identifications were compromised most when a formant synthesis method was employed. As one possible reason for subjects' compromised intelligibility judgements under these circumstances Hillenbrand and colleagues suggested that this method led to a decrease of the originally present spectral cues. Sinusoidally synthesised signals, however, were less likely to be conflated with adjacent vowels with shorter typical durations, presumably because this synthesis technique preserved the spectral cues of the original vowels more accurately.

The pattern emerging from these data is the following: when spectral cues are sufficient, adult English-speaking listeners appear to make use of them; but a distortion of spectral cues seems to induce listeners to turn to duration cues, perhaps in an attempt to compensate for the distortion of the signal. Hence a change in the context or in the reliability of the primary cues may lead to a shift in subjects' attention to other accompanying acoustic variation. Hirata (2005) found that English listeners are able to learn to make some use of vowel duration cues for the purpose of vowel identification in non-native listening, while also demonstrating less perceptual flexibility than native listeners.

Hirata's study explored two different factors that may pose a challenge for native English speakers' identification of Japanese vowel length, namely speaking rate and pitch accent. Pertaining to speaking rate, it was found that native English listeners' identifications of vowel length were affected by the vowels' absolute durations (which obviously differ at variable speaking rates). Hence, while native Japanese listeners tended to make use of automatic rate normalization when deciding whether a vowel should be categorized as either long or short, non-native English listeners appeared unable to do so. Instead, English listeners only classified Japanese long vowels spoken at slow rates as long whereas Japanese long vowels spoken at fast rates were categorized as short. The slow rate obviously made these vowels 'extra-long' while the fast rate reduced the vowels' absolute durations.

Overall, the impression emerges that but the role of vowel length and the processing of its variation differ across English, Japanese and Dutch. Most importantly for the present study, English native listeners do not tend to use vowel duration as a primary phonemic cue for vowel identification in the same way as Japanese or Dutch listeners do, although English listeners do appear to be sensitive to this information and make some use of it for other aspects of processing (e.g. for determining the voicing status of postvocalic obstruents and fricatives) and under certain circumstances (e.g. when spectral cues are insufficient or have been distorted).

English listeners' sensitivity to vowel length is likely to be the result of two factors: first, vowel length undoubtedly has phonological relevance in English, as previously discussed. Moreover, vowel length may be salient acoustically (see section 2.6.1) and hence more likely to be perceived by listeners, even if the same listeners may not make spontaneous use of it for determining vowel identity under more natural conditions and when adequate spectral cues are present in the incoming signal.

When the role of vowel duration in Japanese and English is compared with its role in Dutch, the latter may be argued to take an intermediate position.

Although there is a phonemic vowel duration contrast in Dutch (Nooteboom, 1972) it does not affect all vowel types equally and in the same way. For the vowel pairs /a:, ɑ/, /o:, ɔ/, /ø, œ/ and /e, ɪ/, duration appears to be relatively more important than vowel quality for Dutch listeners to determining the vowels' identity (Nooteboom, 1972). Hence, when discriminating minimally different Dutch words containing the vowels /ɑ/ versus /a:/, as in 'tak' (i.e branch) versus 'taak' (i.e. task), native listeners use vowel duration information to tell the vowel and thus the words apart (Nooteboom & Doodeman, 1980).

Experimental support for this claim was provided in a large-scale study exploring listeners' knowledge of acoustic phonetic information during segment identification (Smits, Warner, McQueen & Cutler, 2003). Using a gating technique, Smits and colleagues demonstrated that shortening the long Dutch /a/ such that its length resembled that of the short Dutch /ɑ/ resulted in many listeners labelling the long /a/ as its short counterpart /ɑ/. This tendency was much weaker for a Dutch /i/ fragment that had the length of its short spectrally more dissimilar counterpart /ɪ/. This differential categorization pattern for the two types of vowel pairs suggests that vowel quality may have been more important perceptually in the latter case (i.e. for /i/ versus /ɪ/), whereas vowel length was more crucial for distinguishing the former pair (i.e. /ɑ/-/a/). Importantly, even for the vowel pairs that do appear to be distinguished based on vowel length cues, there are accompanying spectral differences and hence vowel quality differences. This issue will be discussed further below in section 3.2.

Dutch vowels that do not appear to participate in a quantity opposition are the three vowels /i/, /y/ and /u/ (Nooteboom, 1972; p.26). Nooteboom pointed out that these 'exception' vowels behave as phonetically short before /p/ and /t/ and as long before /r/ (Nooteboom, 1972; p112).

Learning about the specific role that vowel length plays in one's native language is undoubtedly complicated for the language learner by the fact that vowel length also varies as a function of context (see Erickson, 2000 for a review). Among the many contextual factors that influence vowel length to differing extents across many languages are the aforementioned voicing status of the following consonant (e.g. House, 1961; Luce & Charles-Luce, 1985; also see Watson, 1983, for a review) and the number of syllables in the word (e.g. Nooteboom, 1972). For example, vowel length is reduced as the number of syllables in the word increases (the latter appears to be particularly true for stressed vowels) (Nooteboom, 1972). Lexical and emphatic stress (Lieberman, 1960; Botinis, Banner, Fourakis & Pagoni-Tetlow, 2002; see Crystal & House (1988), for a review), the position of the word within the phrase (Oller, 1973), and speaking rate (Hirata, 2004; Botinis et al., 2002) have also been found to influence the duration of vowels.

Given this substantial variation and the likely covariance of the above factors in continuous speech (Erickson, 2000), it is likely to be a considerable challenge for infants to discover the role of vowel length in their native language and the latter makes our investigation an intriguing area of research.

3.2 Experiment 2

In the present study we investigated whether Dutch infants would consider the vowel length opposition important in a word categorization task using the Conditioned Headturn method.

As noted, the Dutch vowel inventory consists of long and short vowels, some of which convey contrastive meanings (Nooteboom, 1972), although there is an accompanying vowel quality difference even between those long and short Dutch vowels that convey contrastive meanings (Nooteboom et al., 1980). For example, the Dutch long /a/ is more open and higher in vowel space than its short counterpart /ɑ/. However, for this vowel pair the durational cue seems to take precedence, as described above.

Experiment 2 involved Dutch-learning infants being trained to make a headturn response to a Dutch word form containing the short Dutch vowel /ɑ/. During testing we measured infants' responses to the target word form ('tak') and to two variant word forms differing either in vowel length ('taak') or vowel quality ('tek'). The test words were again embedded in utterances just as in Experiment 1.

3.3 Method

3.3.1 Participants

Our subjects were 20 monolingual Dutch infants ranging from seven and a half to twelve months (mean age: 288 days). Participating subjects were assigned to two groups (Vowel Quality versus Vowel Length Group). Each group consisted of 10 subjects. Nine additional infants were tested but not included in the final sample because of crying, fussiness, failure to meet the training criterion, the parent's interference during testing or experimenter error. Participants were recruited using the same criteria as in Experiment 1; exclusion criteria were also the same (see section 2.3.1).

3.3.2 Materials

The stimuli were comparable to those in Experiment 1. They consisted of five isolated words (including the target) and a range of short simple utterances. A list of the sentences and of the background words is provided in the Appendices (section 7.2). The stimuli were spoken by an female native speaker of Dutch. The target word itself was changed from 'hasp' in the first experiment to 'tak' in the following two experiments. The test sentences and the variant word forms are also listed in the Appendix to this chapter. The variant words were the length variant 'taak' and the quality variant 'tek'.

The rationale for the aforementioned change of the target word from 'hasp' to 'tak' was as follows:

First, we hoped that this change to the stimuli would make the task easier for infants, thereby raising the number of correct headturns overall and increasing the stability of the proportional response measure. Our reasoning for the task being more simple was as follows: /tɑk/ contains no clusters and it is thus not as complex a word as 'hasp'. This may make the word form 'tak' easier to process and retain. Furthermore, its consonants are obstruents, which may be perceived more readily than the low energy uvular fricative /h/ in 'hasp'. The second reason for the change of the stimuli was that we intended to manipulate the vowel in 'tak' to isolate the vowel duration cue; manipulating a vowel is easier when it is surrounded by obstruents rather than fricatives.

As in the first experiment, there was a phonemic resemblance between the background words and the target word to increase infants' attention to the vocalic portion of the target. Target words were always in sentence-final position. Words commonly known to infants at their age as well as the target vowel /ɑ/ were avoided in the carrier sentences for the same reasons as outlined in section 2.3.2. Unlike in Experiment 1, there were three different tokens of each sentence type in the present experiment. This is to say, each carrier sentence, ending in 'tak', 'taak' and 'tek' (e.g. 'Peter tekent een tak/taak/tek'), was presented three times. This obviously reduced the total number of sentence types and may therefore have made the experiment less challenging for infants. At the same time, the fact that infants heard the same carrier sentence ending in either the trained target or its variants could also be misleading and increase the proportion of false positives.

The 'taak' tokens, which served as the vowel duration variant in the present experiment, were created by lengthening natural tokens of 'tak', using PRAAT sound-editing software (Boersma & Weenink, 2002). A summary of the key steps of the manipulation procedure is provided in the Appendices (section 7.4).

As we were interested in duration per se, and for comparability with the English experiment (Experiment 3), only duration was manipulated in Experiment 2, even though this created a non-native Dutch vowel. The vowel was non-native because, strictly speaking, there is a length difference and a quality difference between the Dutch vowel /ɑ/ and its long counterpart, as previously noted. Hence the phonetic symbols /ɑ/ and /a:/ will be used to refer to the Dutch phonemes and the symbol /ɑ:/ will be used to refer to the artificially lengthened sound /ɑ/ in the present experiment.

3.3.3 Procedure

As the procedure was described in considerable detail in section 2.3.3.1, mainly aspects pertaining specifically to the present experiments will be noted here. Experiment 2 involved an exploration of Dutch infants' similarity judgements of word forms in utterances. The word forms differed either in vowel *quality* or in vowel *duration*.

A between-subjects design was used to allow more trials in each condition per infant, in an attempt to enhance our estimate of each subject's mean in each condition.

Training involved infants being conditioned to respond to isolated tokens of the word form /tɑk/ containing the target vowel /ɑ/. Infants were subsequently tested on word forms embedded in utterances. The vowels of the embedded target words varied. Generalization was tested using a lengthened version of the native vowel /ɑ/ (i.e. /ɑ:/). We also incorporated a control condition, namely the qualitatively different vowel /ɛ/. This is to say, sentences ended in either 'tak' 'taak' or 'tek'. We selected the control vowel /ɛ/ because we planned to test Canadian babies in

Experiment 3; Dutch and Canadian /ɛ/ are very similar. Hence selecting /ɛ/ allowed us to use the same comparison vowel in both subject groups.

The present experiment consisted of two between-subjects conditions involving the word forms 'tak-taak' (i.e. the Vowel Length/Duration Condition) and 'tak-tek' (i.e. the Vowel Quality Condition) (see Figure 3.1). The latter served as our baseline or control.

Training involved infants being conditioned to turn their head towards a visual reinforcer when they heard the novel word form 'tak'. As in Experiment 1, infants were presented with background word forms such as 'tok' and 'tik'. Whenever the training word 'tak' was presented the visual reinforcer was activated. The variant word forms (here: 'taak' and 'tek') were again not among the background tokens.

Infants were trained to turn their head when they heard a word containing the vowel /ɑ/

e.g. tik tok ket /tɑk/ /tɑk/

↓ ↓ ↓ ↓ ↓

background words target word

headturn was trained here

Figure 3.1: Training phase

During the test phase infants heard utterances containing either the trained target 'tak' or the variants 'taak' and 'tek', embedded in utterance-final position. Table 3.2 presents an example of a Dutch test sentence and its three embedded variant word forms. In this example, babies in the Vowel Quality Condition heard either 'Rian beschrijft een tak' (i.e. Rian describes a 'tak') or 'Rian beschrijft een tek' and babies in the Vowel Duration Condition heard either 'Rian beschrijft een tak' or 'Rian beschrijft een taak'. We measured proportions of headturns across conditions.

Example sentence during test phase of CHT-study	
Rian beschrijft een /tɑk/	Trained Target
Rian beschrijft een /tɑ:k/	Vowel Length Condition (native)
Rian beschrijft een /tɛk/	Vowel Quality Condition (native)

Table 3.2 Conditions during test phase

3.4 Predictions

As can be seen in Table 3.3, we predicted most headturn responses to the trained target. We also expected that false alarms to 'tek' and to 'taak' would be equally low, in line with Dutch segmental phonology.

Predictions for Dutch infants			
Hits	tak	>	tek & taak
False alarms	tek	=	taak
• Categorization in line with Dutch segmental phonology			

Table 3.3: Predictions for Experiment 1

These predictions were based on our hypothesis that infants would categorize /ɑ/ and /ɑ:/ as well as /ɑ/ and /ɛ/ as different vowels due to the phonemic structure of Dutch (i.e. both /ɑ/ and /ɑ:/ and /ɑ/ and /ɛ/ are contrastive in Dutch). Consequently, 'tak' versus 'taak' and 'tak' versus 'tek' should qualify as possible Dutch words and hence mark lexical distinctions for infants.

If the above predictions were borne out this would suggest that both vowel quality and vowel length signal a lexical distinction for Dutch infants in this experiment. However, as the Dutch target vowels differing in vowel duration (i.e. /ɑ/ and /ɑ:/) are indeed short and long vowels of Dutch the predicted findings would not serve as evidence that Dutch infants attend to vowel length differences when listening to Dutch vowels in general, let alone when listening to non-native vowel sound contrasts.

3.5 Results

The primary question to be addressed was whether Dutch infants would respond most to the trained target and relatively little to both variant word forms since both differ phonemically from the trained target.

Figure 3.4 presents the proportion of HT-responses across conditions for infants in the Vowel Quality Group and Figure 3.5 for infants in the Vowel Length Group. As predicted, infants in the 'tak-tek' condition turned most frequently to the training word 'tak' and much less to 'tek'. The same was true for babies in the 'tak-taak' condition.

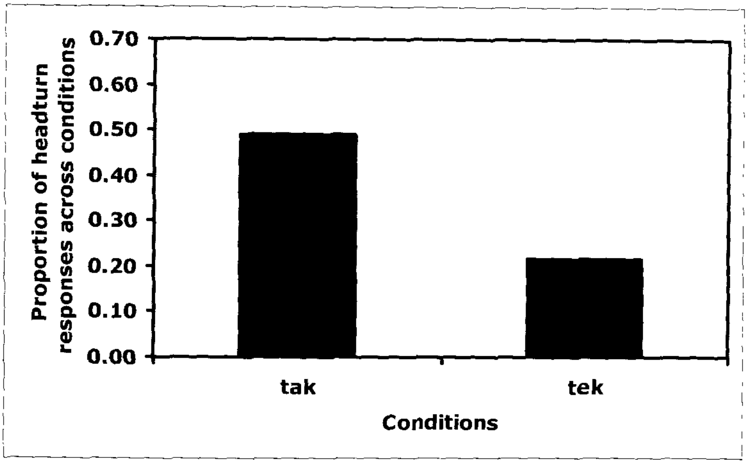

Figure 3.4: **Dutch Vowel Quality Group**

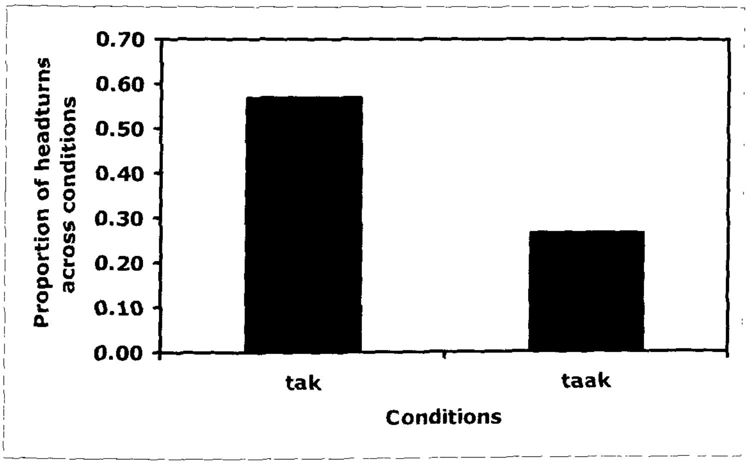

Figure 3.5: **Dutch Vowel Length Group**

Paired samples t-tests were conducted over proportional measures of hits versus false positives for each of the two groups (i.e. Vowel Quality and Vowel Length) to examine whether Dutch infants in both groups turned more readily to the trained target as opposed to its variants. These analyses revealed significant differences between proportions of headturn responses both for the Vowel

Quality Group ($t(9)= 4.18$, $p< .02$, all tests two-tailed) and for the Vowel Length Group ($t(9)= 4.81$, $p< .01$).

Furthermore, a One-Way ANOVA was conducted to explore whether the significant differences in the proportions of headturn responses to the target and its variants reported for the Vowel Length as well as the Vowel Quality Group were the same. Both hits and false positives were compared across the two groups. Response proportions for hits did not differ across groups ($F(1, 9)= 3.03$, n. s.), nor did response proportions for false alarms ($F(1, 9)= .64$, n. s.).

In summary, Dutch infants in the Vowel Length and the Vowel Quality Group responded most to the trained target word 'tak' and much less to its variant word form 'taak' and 'tek' respectively. Moreover, a comparison of the differences in the proportions of headturn responses to the target and its two variants across the two groups revealed that they were not significantly different.

3.6 Conclusions

In the present experiment Dutch infants tended to categorize both variant word forms (i.e. 'taak' and 'tek') as lexically distinct from the phonemically different trained target 'tak'. These findings are in line with our hypothesis that infants' generalization pattern would be influenced by Dutch segmental phonology. Moreover, the findings are consistent with infants' performance in Experiment 1 (see section 2.5): in both cases Dutch infants showed evidence of perceiving the vowel length variant of the target word and the native vowel quality variant of the target word as different words from the trained target, although the specific word forms employed differed across the two experiments.

Although suggestive, these findings in isolation are not sufficient as evidence that it was indeed the segmental structure of Dutch that was responsible for the pattern of results obtained. Another possibility is that the vowel length contrast, being temporal in nature and perhaps acoustically salient in general, may have been noticed by infants, irrespective of the native language segmental phonology.

Experiment 3, therefore, served to help us interpret Dutch infants' response pattern and to address the aforementioned question of whether the native language phonology affects infants' perception of vowels differing in length during word learning: here, we tested infants learning Canadian English, which does not have a phonemic vowel length contrast on similar stimulus materials.

3.7 Experiment 3

As outlined in the introduction, Dutch and English have different vowel inventories and vowel length plays a different role in these two languages. For some Dutch vowels, vowel length is a main cue differentiating two vowels. In English, however, vowel length differences exist in everyday speech but they do not signal a primary difference in vowel identity. Yet English vowel length does serve as a cue to the voicing status of postvocalic stops and fricatives (Denes, 1955, Raphael, 1972; see Watson, 1983, for a review). Specifically, vowels preceding voiced consonants may be twice the length of vowels preceding voiceless consonants (Watson, 1983), though the magnitude of the effect depends on the materials examined (Crystal & House, 1988). English listeners thus do need to pay attention to vowel length when identifying words such as 'tap' versus 'tab'. It is important to note, however, that even in these cases a difference in vowel length alone is unlikely to induce a change in the perception of the voicing status of the following obstruent (e.g. Wardrip-Fruin, 1982; O'Kaine, 1978; Hogan & Rozsypal, 1980; Raphael, 1981; Revoile et al., 1982; Hillenbrand et al., 1984; cited in Hillenbrand et al., 2000).

Given the difference between the role that vowel length plays in Dutch and English for the purpose of vowel identification, it is conceivable that Dutch- and English-learning infants' processing of vowel duration will reflect this different pattern: English infants may pay relatively less attention to vowel length information in vowel identification than Dutch-learning infants when there are no accompanying vowel quality differences and providing the postvocalic consonants are identical. The latter should prevent infants attributing a difference in vowel length as being a cue indicating a voicing difference of the following consonant.

Experiment 3 thus served to explore the question of whether Canadian infants' pattern of generalization would be in line with Canadian segmental phonology.

3.8 Method

The present experiment was conducted in the Infant Studies Centre of the Psychology Department at the University of British Columbia in collaboration with Dr. Janet Werker.

The Exeter Conditioned Headturn Program was employed for this study since the Nijmegen setup could not be transferred to the laboratory in Vancouver. Using a different program entailed some methodological variation. However, we tried to keep the latter to a minimum.

The main methodological and procedural features of the Exeter and hence the Vancouver setup that made it different from the Nijmegen set-up are summarized in Table 3.6. The points will be outlined in more detail in section 3.8.3.

Summary of main features of Exeter Conditioned Headturn Program
Two-person procedure (experimenter and assistant)
Online randomization of stimuli
Additional conditioning phase following the initial training phase–this additional phase involved reinforcement contingent on the infants' headturn behaviour as well as allowing re-training trials in cases where infants had had a pre-specified number of misses.
Timing of gradual delay of reinforcement controlled internally and hence not adjustable by specifying time in ms post word-onset .

Table 3.6: Exeter Conditioned Headturn Program

3.8.1 Participants

Our subjects were 32 Canadian infants in the age-range of seven and a half to twelve months (mean age: 320.61 days). Participating subjects were again assigned to two groups: Vowel Quality (n=16) versus Vowel Length (n=16). 14 additional infants were tested but not included in the final sample because of crying, fussiness, failure to meet the training criterion, the parent's interference during testing (e.g. taking off of headphones), experimenter- or equipment error.

Most infants were recruited by visiting new mothers at the BC Women's Hospital in Vancouver. During that visit consent was obtained to contact parents by telephone to invite them to participate in specific studies when their infants had reached the appropriate age. Other parents had expressed their interest at public service announcements or by responding to advertisements in local newspapers. All infants were at least 38 weeks gestation, had normal birth records and no history of health problems. Exposure to English was at least 70% by parental estimate. When infants had had exposure to a second language it was ensured that this was not a language with phonemic vowel length.

3.8.2 Materials

The stimuli were comparable to those in Experiment 1 and 2, consisting of five isolated words and a range of short simple utterances. The stimuli were spoken by a female native speaker of Canadian English. The isolated words that were used as background stimuli during the training phase again had phonemic resemblance with the target word, which in turn was phonemically similar to the target in Experiment 2. The main difference between the background- and the target word forms was that they were anglicized for the present experiment such that they would indeed be a possible word form for Canadian English infants (i.e. the target word form was changed from /tɑk/ to /tæk/).

3.8.3 Procedure

Unlike the Nijmegen paradigm, the present experiment involved an experimenter and an assistant.

The experimenter's role was to start trials and to indicate whenever the infant had made a headturn response by pressing a button on a joystick. The assistant sat opposite the parent and the child in the experimental room and distracted the infant with silent toys (see 2.3).

Trial orders were randomized online. The initial training phase was shorter than that of Experiment 2 but otherwise identical; a shorter training phase was used to compensate for the addition of a conditioning phase, which was an extension of the initial training. It was felt that the addition of a conditioning phase may reduce the attrition rate in providing infants with more focused teaching: In the conditioning phase, infants were only reinforced for correct headturns, rather than for all trials on which the target was presented. Infants progressed to the test phase after meeting a criterion of five out of six correct responses (including correct rejections) in the conditioning phase. The total number of trials in both setups averaged 30 trials, although the exact number was determined by each infant's performance.

The design was a between-subjects design, as in Experiment 2. Infants were trained to respond to the word form 'tack' (pronounced /tæk/) containing the Canadian vowel /æ/. Generalization was then tested on one of two variant word forms containing different vowels. Analogous to those in Experiment 2, the embedded vowels were a lengthened version of the training vowel /æ/ (i.e. /æ:/ as in 'taack') and the qualitatively different variant /ɛ/ ('teck'). Canadian infants were tested in the same way as Dutch infants. The main difference between the Dutch and the Canadian CHT-study was that we obviously had to use Canadian English stimulus materials for reasons of consistency across the two experiments. The background words were kept as similar as possible, although individual sounds were anglicized where necessary.

3.9 Predictions

If Canadian infants ignored vowel duration in their head-turning responses, it would suggest that they had learned that vowel duration per se is not an important component of a vowel's identity and a word's phonology. If Canadian infants responded as the Dutch infants did, by failing to generalize to lengthened variants of the target, it would suggest one of the following: either that these infants had not learned to ignore vowel length variation, or that they were focusing on vowel length as a cue to voicing in the coda consonant. The latter is less likely though, given that there were no accompanying cues inducing infants to make this interpretation and considering the findings referred to by Hillenbrand et al. (2000) mentioned in section 3.1.

As presented in Table 3.7, our predictions for the qualitatively different word forms were the same as those for Dutch infants, namely that infants would not readily generalize from 'tack' to its quality variant 'teck'. However, predictions for the quality variant 'taack' differed: Based on the different nature of the segmental phonology of Canadian English we hypothesised that Canadian babies would regard vowel length as relatively less important compared with Dutch infants and that they would be unlikely to regard the word forms only differing in vowel duration as two different word forms. Accordingly, we expected more false alarms to the lengthened variant /tæ:k/ than to phonemically different word form /tɛk/ due to its conflation with 'tack'.

Canadian infants were thus predicted to generalize more readily from /æ/ to /æ:/ than from /æ/ to /ɛ/, providing their native language phonology dictates their pattern of responses in this task.

However, given the possibility of vowel length being relatively more salient than non-native vowel quality contrasts, it was considered possible that infants would still be able to discriminate the word forms differing in vowel length. Nevertheless, we did not expect this ability to be reflected in our categorization results.

Predictions for Canadian infants			
Hits	tack	>	teck
Hits	tack	=	taack
False alarms	teck	<	taack
• Categorization in line with Canadian segmental phonology			

Table 3.7: Predictions for Experiment 3

3.10 Results

The central issue to be addressed here was whether Canadian infants would show a different pattern of responses compared with Dutch infants based on differences of Canadian segmental phonology. Specifically, a similar response rate was expected to the trained target and its length variant but fewer responses were predicted to the quality variant for reasons discussed in the previous section.

Figure 3.8 presents the proportion of HT-responses across conditions for infants in the Vowel Quality Group and Figure 3.9 for infants in the Vowel Length Group. Infants in the 'tack'-'teck' condition turned considerably more to the training word 'tack' and much less to 'teck'. Surprisingly, the same was true for babies in the 'tack'-'taack' condition.

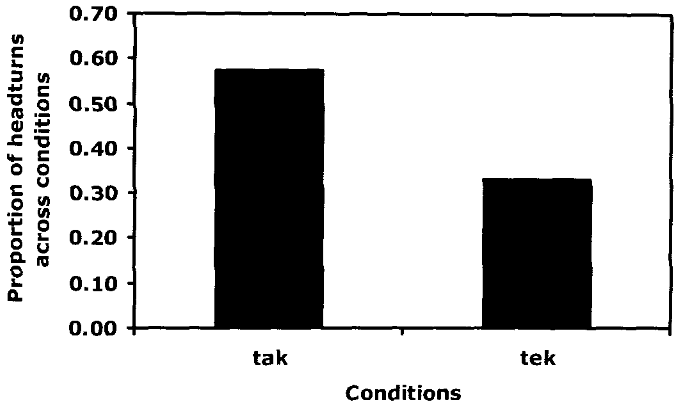

Figure 3.8: **Canadian Vowel Quality Group**

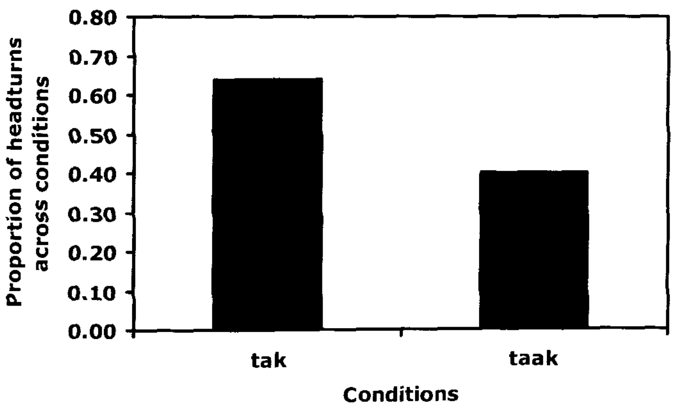

Figure 3.9: **Canadian Vowel Length Group**

Paired samples t-tests were conducted over proportional measures of hits versus false positives for each of the two groups (i.e. Vowel Quality and Vowel Length) to examine whether Canadian infants turned more readily to the trained target as opposed to one or both of its two variants. These analyses revealed significant differences between proportions of headturn responses both for the Vowel Quality Group ($t(15)= 3.58$, $p< .003$, all tests two-tailed) and for the Vowel Length Group ($t(15)= 3.11$, $p< 0.007$).

Just as in Experiment 2, we also conducted a One-Way ANOVA comparing differences in proportions of headturn responses to the target and its variants across the two groups (i.e. Vowel Length and Vowel Quality).

Again, hits and false positives were compared across groups. In line with our predictions, response proportions for hits did not differ across groups ($F(1, 15) = .62$, n. s.). Surprisingly, response proportions for false alarms were also the same across the Vowel Length- and the Vowel Quality Group ($F(1, 15) = .34$, n. s.). The latter finding was contrary to our predictions.

In summary, Canadian infants responded most to the trained target word 'tack' and much less to both its variant word forms 'teck' and 'taack'. Moreover, there was no difference in the effect size across groups.

The present findings thus match the pattern of results obtained for Dutch infants in Experiment 2.

3.11 Discussion

Overall, Canadian English infants' pattern of responses indicates that they generalized most to the trained target word form and significantly less to both variant word forms. Contrary to our predictions, there was no difference between infants' generalizations (i.e. false alarms) to the two variant words 'teck' and 'taack' in the present experiment.

Taken together, these word form generalization findings from Dutch and Canadian babies in Experiment 2 and 3 suggest that both groups of infants seem to use vowel length cues when categorizing word forms in utterances. As noted, the Canadian pattern of results is surprising given that it does not conform to Canadian segmental phonology.

How can Canadian infants' sensitivity to vowel length in this task be explained? There are at least two possible explanations for this pattern of responses:

First, as was pointed out above, vowel length, although not a phonemic cue to vowel identity in English, cannot be ignored by English learning infants because it can serve as a important contributing cue for other aspects of processing in speech perception; among these are the determination of the voicing status of the following coda consonant in words such as 'tap' and 'tab'. English learning infants thus face the challenge of discovering the importance of vowel length at some but not all levels of analysis in speech perception. While this may be true for many perceptual features that infants become attuned to during language acquisition, vowel length may be 'special' in the sense that it is salient acoustically, at least in the present series of experiments and perhaps even generally so (i.e. see earlier discussion of the use of vowel length during L2-listening). As a result, vowel length may affect Canadian infants' categorization pattern even in cases where the native language segmental phonology makes a different prediction. Supportive evidence for even newborns' sensitivity to vowel duration cues comes from a study by Christophe, et al., 1993). Christophe and colleagues demonstrated that three-day-old infants were able to discriminate two sets of phonemically identical CVCV disyllables that had been extracted either from between two words (e.g., 'mati' in panorama typique) or from within a word (e.g., mati in mathématicien). Infants were tested using the high-amplitude non-nutritive sucking procedure (described in detail by Jusczyk, 1985). As there were consistent differences in the duration of the vowel /ɑ/ between the two categories (i.e., /ɑ/ was on average 15 ms longer in the Between Word Condition compared with the Within Word Condition), it is conceivable that even in newborns such small differences in duration may serve as a potential word-boundary cue.

Second, it is possible that the CHT-paradigm is simply not sensitive enough to pick up existing language-specific differences in processing between Dutch and Canadian infants. Specifically, the CHT-task employed merely requires infants to compare a taught word form with a set of similar sounding comparison word forms. While this demands discrimination of small phonetic distinctions, it may not encourage infants to draw upon their native language sound categories in that no mapping of word forms onto objects or pictures of objects is necessary. As a result, in spite of our having placed the target words in sentences, it is conceivable that infants' processing of the stimuli was 'shallow' – more phonetic than phonological.

Therefore, an obvious but very important issue to address is what our task really measured. There are at least two possibilities: As discussed above, it may have measured no more than raw similarity independently of the phonology. Specifically, 'tack' and 'taack' could be phonologically equivalent for Canadian infants but because 'taack' sounds different from 'tack' at an acoustic level, infants may have responded less to it. Secondly, it is also possible that the task did indicate a phonological difference between 'tack' and 'taack'. If English infants did perceive the word forms as segmentally different, they would clearly need to learn to suppress this segmental interpretation eventually for the purpose of vowel identification.

On either account, resolving this issue may require the use of a more referential task. This was done in the experiments that follow. If Canadian infants continued to discriminate word forms differing only in vowel length, this would corroborate the hypothesis that Canadian infants are indeed perceiving vowel length differences at a segmental level. The latter appears unlikely, however, given the segmental phonology of Canadian English. It is more likely that a task tapping a lexical level of processing would result in infants ignoring vowel length differences, at least eventually. It may still be the case that the acoustic salience of vowel length results in infants' attunement to that cue being delayed relative to other non-native phonetic cues (e.g., the retroflex-dental contrast studied in detail by Werker and Tees (1984) and Werker and Lalonde (1988)).

Previous infant perceptual experiments have largely focussed on maintenance or l oss of very subtle distinctions. Here, we are considering a different problem, namely how infants learn to assign salient variation to the correct part of their phonology. Understanding the answer to this question will help us determine how infants' perceptual sensitivities develop into language-specific phonological knowledge.

In the following experiments we employed a word-learning task to test whether 17-18-month old Canadian- and Dutch- learning children are equally capable of learning novel word forms that differ in vowel length. For comparison, two word forms differing in vowel quality (i.e. a phonemic difference for both groups of infants) were also included as a control measure to help us evaluate the magnitude of the effects of the duration manipulation relative to a quality manipulation.

4 VOWEL DURATION IN A REFERENCE ASSIGNMENT TASK

4.1 Introduction

Experiments 4 and 5 were motivated by our findings that Canadian infants' low-level similarity judgments of minimal-pair words differing only in vowel length could not be predicted solely on the basis of their native segmental phonology.

It is unlikely that the lack of language-specificity in our data is exclusively the result of the experimental paradigm (i.e. the CHT-method) not being sensitive enough in general to reveal language-specific influences. First, we did obtain positive evidence of discrimination, albeit not language-specific in nature. Second, Werker and Tees employed the CHT-procedure in their aforementioned attunement studies, which revealed language-specific effects, as have other studies based on this method. The procedure may still play a role, however, in terms of its processing demands in that it involves low-level and hence non-lexical processing, as discussed in section 3.11. Moreover, if it is indeed the case that vowel length is acoustically more salient relative to other non-native phonetic differences that have been tested to date using this method, one may still find language-specific differences in performance between the two language groups when using a different method, tapping into a more lexical level of processing. This possibility was explored further in the present experiment.

The acoustic salience of vowel length cues may pose a particular challenge for infants in that their salience could make stimuli of differing vowel length 'pop out' perceptually. This may then lead to vowel length cues being more difficult to 'ignore' for those aspects of processing for which they are not only not essential but may even be misleading. An example of such a case was discussed earlier: Japanese L2-listeners' use of vowel length when categorizing the Canadian English vowels /i/ and /ɪ/ and the resulting misperception of these vowels in some phonotactic environments (Morrison, 2002). The relevance of these data for the present argument is limited by the fact that Japanese listeners appeared to be drawing upon their native language vowel duration categories in this task (Morrison, 2002); thus the data cannot provide proof that non-native listeners for whom duration cues are not meaningful for vowel identification in their native language would also draw upon this source of phonetic variation for the purpose of vowel categorization. To explore this question, Morrison also tested Spanish listeners, for whom duration plays no role in vowel identification. Morrison found that Spanish listeners appeared to make use of both durational and spectral information, although overall Spanish listeners' categorizations appeared to be more random (Morrison, 2002). Hence vowel duration may or may not turn out to be salient perceptually for non-native listeners.

For English-learning infants, vowel length will clearly be helpful for determining the voicing status of post-vocalic plosives, but attending to this source of phonetic variation does not appear as helpful as a primary cue for discriminating words such as 'hit' and 'heat'. Although the tense vowel in 'heat' does tend to be longer than the lax vowel in 'hit' (providing contextual factors such as speaking rate remain constant), adult native English listeners attend primarily to spectral information when categorizing these two vowels, as noted earlier.

In Experiment 4, it was of primary interest to us whether Canadian English children would be able to learn that two word forms, differing solely in vowel length, can refer to two different objects. The paradigm employed was the Switch Paradigm (Cohen, 1992) since it draws upon more referential learning than the CHT-method.

To ascertain whether our results in Experiment 4 can be interpreted as the reflection of a cross-linguistic difference between Dutch and Canadian infants, induced by the segmental phonology of the two languages, the same study was carried out with Dutch children in Experiment 5.

It is also important to note that subjects in Experiment 4 and 5 were older than the infants we tested with the CHT-paradigm (i.e. they were on average 18, as opposed to 10 months old). Consequently, participating subjects in the present two experiments have had considerably more exposure and hence listening experience with their native language, Canadian English and Dutch respectively. As the distribution of vowel length in English infant-directed speech is likely to be unimodal (Werker, Pons, Dietrich, Kajikawa, Fais, & Amano, submitted) rather than bimodal, participants in Experiment 4 are unlikely to have set up a category difference for spectrally identical long and short vowels of the ambient language (Maye, Gerken & Werker, 2002). The study by Maye et al. showed that exposure to an artificially created continuum of consonants whose length was bimodally distributed but not exposure to consonants with a unimodal distribution of length lead to infants establishing two distinct sound categories during a subsequent discrimination task.

4.2 Experiment 4

Experiment 4 was an attempt to explore whether Canadian infants would still consider the vowel length opposition important when they were engaged in an activity tapping into a more referential level of processing.

4.3 Method

We used the Switch method for the present experiment as well as for all experiments that follow (see section 3.3.3).

Infants were habituated to two novel word-object pairings and subsequently tested on both a familiar and a 'switched' pairing to ascertain whether they would detect the violation (details about the procedure will be discussed in section 4.3.3).

The setup was comparable to the one described in detail by Werker & Fennell (2002). The experiment took place in a quiet, dimly lit room (approximately 6.5 m^2 in size). The infant was seated on the parent's lap facing a television monitor located 1.5 metres away from the infant. The monitor was surrounded by black cloth, which also concealed a Sony digital video camera located underneath the monitor. The camera served to record infants' eye-movements for online coding from a control room. The loudspeaker delivering the speech stimuli was located above the monitor. A Macintosh G4 was used to set up and run the program. All visual and auditory stimuli were directed from the computer to the testing room. Auditory stimuli were presented to infants at a comfortable loudness level (approximately 68-72 dB SPL). The experimenter was seated in a control room from which she pressed a designated key on a computer keyboard whenever and for as long as the infant looked at the television screen during trials. It was also possible to start new trials by pressing a designated computer key when the infant was oriented towards the television screen. Both the experimenter and the parent were blind to the stimulus being presented and hence to the condition and the experimental phase.

4.3.1 Participants

Our subjects were 33 Canadian infants ranging from 16 months 16 days to 20 months, 19 days (mean age: 18 months, 20 days). Participating subjects were assigned to two groups: Vowel Quality (n=18) and Vowel Length (n=15). 17 additional infants were tested but not included in the final sample because of crying, fussiness, failure to habituate, parental interference during testing, experimenter/equipment error or sibling's reluctance to allow parent to enter testroom with participating child. Infants were recruited as described previously, using the same selection criteria.

4.3.2 Materials

Our speaker was a female native speaker of Canadian English. The stimuli consisted of 20 isolated CVC-word forms of four types, 'tam', 'tem', 'taam' and 'neem'; there were five tokens of each type. The 'taam' tokens were created by lengthening (i.e. doubling) natural productions of 'tam'. Manipulations were done in the same way as in the previous experiments.

Average spectral and duration measurements for the 'tam', 'tem' 'taam' and 'neem' tokens used for the present study are displayed in Figure 4.1.

Canadian Stimuli	Mean word duration	Mean vowel duration	Mean F1	Mean F2
tam	609.96	298.87	695.97	2127.11
tem	541.54	205.62	803.06	1943.13
taam	881.04	576.48	695.97	2127.11
neem	736.86	346.86	436.86	1484.44

Figure 4.1: *Acoustic measurements for Canadian stimuli*

As 'taam' was an artificially lengthened version of 'tam', F1 and F2 were identical. 'neem' served as a control word. For a discussion of durational differences between Canadian English versus Dutch stimulus materials, see section 4.8.2.

4.3.3 Procedure

The Switch paradigm is an audiovisual habituation procedure, which was first described by Cohen (1992) and subsequently modified by Stager and Werker in collaboration with Cohen and colleagues at the University of Texas at Austin for use in exploring word-object associations in young infants (Werker, Cohen, Lloyed, Casasola & Stager, 1998). The experimental software employed was Habit 2000. The latter was developed by Cohen and colleagues.

There were two phases, a habituation phase and a test phase. We were interested in whether Canadian English infants would link two novel word forms differing only in vowel length with two novel objects. Given that 17-month-olds, but not 14-month-olds, are able to associate two phonetically similar novel word forms with two novel objects (Werker et al., 1998; Pater, Stager & Werker, 1998) we decided to test 18-month-olds to make sure that task demands would not exceed infants' processing capabilities.

4.3.3.1 Habituation phase

Infants were initially habituated to a visual display of two novel objects, one paired with word A and one paired with word B.

The novel word forms 'tam' and 'taam' were paired with two single animated pictures of novel objects. The 'tam' and 'taam' trials alternated until infants habituated to the pairings. Habituation required their looking time at the display to decrease to a pre-specified criterion level (here 50%), relative to the respective infants' longest total looking time across a two-trial block. This is to say, the exact criterion differed for each infant. Looking time was measured online and the program automatically switched to the test phase of the study when the habituation criterion level had been reached. As mentioned in section 4.3, the experimenter was unaware of when the test phase began; this precaution guarded against observer bias effects. Object movement facilitates performance in this task (Werker et al., 1998).

4.3.3.2 Test phase

When the habituation criterion had been reached, infants were tested on two critical pairings ('switch trials') in which the associative link was violated, i.e. object A was paired with word B and object B with word A. Furthermore, there were two familiar trials ('same trials') during which the familiar pairing was maintained; i.e. object A was paired with word A, just as during the habituation phase. 'Same trials' served as control trials. 'Same'- and 'switch' trials were counterbalanced to avoid order effects.

The rationale of the Switch paradigm is that infants who have made an associative link between the word forms and the objects will notice the switch, leading to longer looking times during 'switch trials' compared with 'same trials' (i.e. dishabituation). Infants' failure to dishabituate does not necessarily imply failure of discrimination, however. Infants may still have learned about the word forms in isolation but failed to make the associative link between the words and the objects (see Werker et al., 1998, Experiment 5 versus 6). If a group of infants were merely discriminating but had not learned the word-object associations, they may show dishabituation if one of the target word forms (e.g. 'tam') were changed to 'cam'. The same group of infants may not, however, notice a switch of the pairings of the familiarized word-object pairings since neither the words nor the objects have changed in the latter case. Dishabituation to the switch of the familiar word-object association implies that infants' ability goes beyond merely discriminating the words and the objects; both word forms must have been associated with the appropriate object and these associations retained in memory – at least for the duration of the study – for the switch to be noticeable.

4.4 Predictions

As presented in Table 4.2 and based on our subjects' native language experience as well as our CHT-findings, we predicted that Canadian infants would indeed notice the switch when the target words differed from each other in vowel quality. The Vowel Length Condition allows us to arrive at a fuller interpretation of the results of the CHT-experiment with Canadian infants, which revealed that they did categorize minimally different word forms differing in vowel length as distinct – contrary to what we had predicted based on the segmental phonology of the language. If Canadian infants do notice the switch in the present experiment then together with our earlier results this suggests that they can use phonetic variation of vowel length for the purpose of vowel identification during referential learning (see section 4.2) of novel word forms. On the other hand, if Canadian infants do not dishabituate to the 'switch' trial in the Vowel Duration Condition in the present experiment this would imply that they do not make use of phonetic vowel length variation during referential learning despite being able to discriminate word forms differing solely in vowel duration at a younger age.

Predictions for Canadian Switch Study			
Vowel Quality	Looking time 'switch' trials	>	Looking time 'same' trials
Vowel Length	Looking time 'switch' trials	=	Looking time 'same' trials
• Categorization in line with Canadian segmental phonology			

Table 4.2: Predictions for Vowel Quality- versus Vowel Length Group

4.5 Results

We were primarily interested in whether Canadian infants would dishabituate to a switch of minimal pairs differing either in vowel quality or in vowel length. Figure 4.3 presents mean looking times across conditions for infants in the Vowel Quality Group and Figure 4.4 for infants in the Vowel Duration Group. Infants in the Vowel Quality Condition looked longer during the 'switch' as opposed to the 'same' trial. There was no significant difference in looking time between the 'same' versus the 'switch' trials for infants in the Vowel Duration Condition.

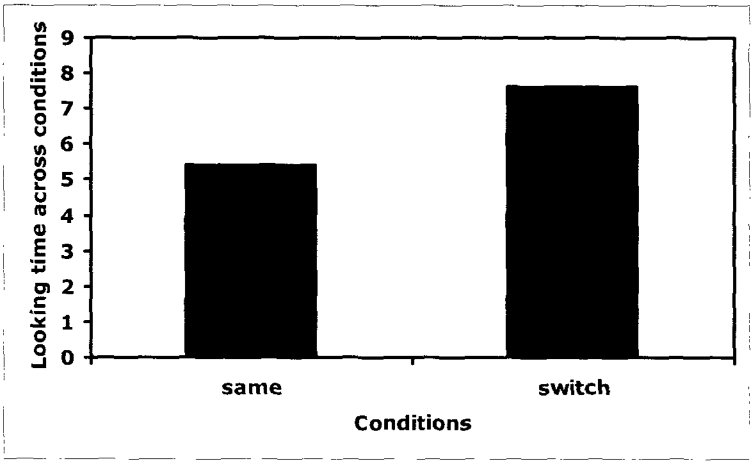

Figure 4.3: **Mean looking times across conditions in the Canadian Vowel Quality Switch Study**

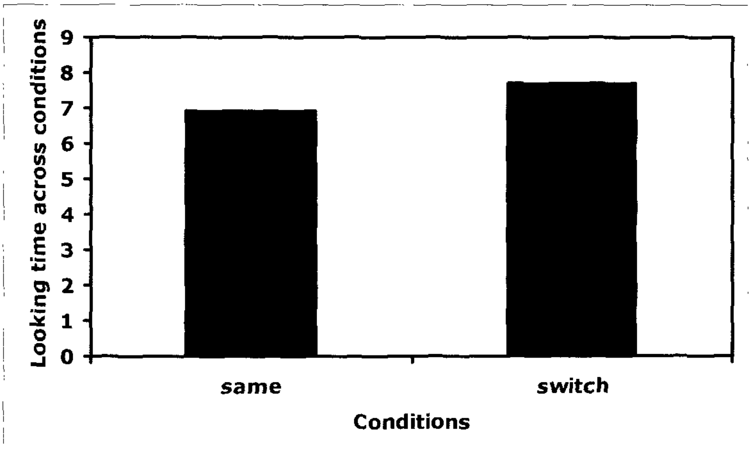

Figure 4.4: **Mean looking times across conditions in the Canadian Vowel Length Switch Study**

First it was determined that infants across the two groups maintained their attention throughout the experiment. This was done by comparing average looking times across the pretest, the last two habituation trials and the post-test, using Planned Contrasts (as in Werker & Fennell, 2004). Importantly, for the Vowel Quality Group, these analyses revealed significant differences between average looking times during the last two habituation trials and the post-test ($t(17)= 12.60$, $p<0.01$, all tests two-tailed), while no difference was present between the pre-and the post-test ($t(17)= 1.62$, n.s.). The same pattern of results emerged for the Vowel Length Group: significant differences were found between looking times during the last two habituation trials and the post-test (t(14)= 12.65, p<0.01), whereas no differences were present between pre-and post-test (t(14)= .27, n.s.).

This recovery of looking time to pretest levels during the post-test across the two groups confirms that the infants were still paying attention during the latter part of the study, hence they increased their attention to the novel post-test stimulus.

The main analyses consisted of Paired Samples t-tests, comparing average looking times during 'same' and 'switch' trials. For the Vowel Quality Group, this analysis yielded significant differences between 'same' and 'switch' trials (t(17)= 2.89, p= .01). However, no differences were present for the Vowel Length Group (t(14)= .54, n.s.).

Summing up, Canadian infants in the Vowel Quality Group dishabituated to the switch of 'tam' versus 'tem', whereas infants in the Vowel Length Group did not dishabituate to the switch of 'tam' versus 'taam' in this task.

4.6 Discussion

Canadian English infants were found to dishabituate to a switch in native vowel quality (i.e. from 'tam' to 'tem'), but not in non-native vowel length (i.e. from 'tam' to 'taam'). Although in line with Canadian segmental phonology, this pattern of results contrasts with our findings from Experiment 3, in which Canadian infants demonstrated discrimination of vowel length in a task involving the CHT-paradigm.

The differential pattern of results in Experiment 3 versus 4 may be due to the nature of the task and the type of processing involved. Whereas the CHT-paradigm in Experiment 3 may have involved processing at a non-lexical level, Experiment 4 was referential in nature thus requiring more high-level lexical processing. Therefore, the latter but not the former could have induced children to focus their attention to phonemically relevant phonetic variation only, thus excluding vowel length.

Importantly, our findings do not preclude the possibility that young children can discriminate the two word forms differing in vowel length for reasons discussed above (i.e. acoustic salience of vowel length and lower task demands), even at this 'advanced' age of 17-18 months and hence long after the internal nature of vowel categories has started to become attuned by the native language vowel inventory.

Could the Switch task simply have been too difficult for Canadian English infants? As discussed in section 4.3.3, Janet Werker and colleagues have shown that 17-month-olds, as well as 14-month-olds tend to dishabituate in the two-picture version of the Switch task when phonemically different word forms are employed: They also demonstrated that 17-month-olds show evidence of dishabituation when phonemically similar novel word forms are used. Our pattern of results, therefore, is unlikely to be due to the complexity of the task. Moreover, if our task had indeed been too difficult, we would also have expected infants to fail to dishabituate in the Vowel Quality Condition, which they did not.

To conclude, the present data are compatible with the hypothesis that Canadian infants have processed the two word form tokens differing in vowel length (i.e. 'tam' and 'taam') in a non-phonemic way, whereas the minimal pair words differing in vowel quality were processed as distinct word forms.

4.7 Experiment 5

Experiment 4, in itself, does not provide sufficient evidence that the pattern of results for Canadian English infants is the result of the segmental phonology of English. Corroborating evidence for this claim would be a different pattern of results for a group of infants for whom vowel length is phonemic in nature. Therefore, Experiment 5 was a replication of Experiment 4 with Dutch children. Due to the differing role vowel length plays in the two languages, a different pattern of performance was predicted for Dutch children in the present experiment.

4.8 Method

The method was the same as in Experiment 4: children were habituated to two novel word-object combinations and subsequently tested on both a familiar and a 'switched' pairing to ascertain whether they would notice either a change in vowel quality (i.e. from 'tam' to 'tem') or in vowel length (i.e. 'tam' to 'tem'); for procedural details, see section 4.3.3.

The Nijmegen (Dutch) and Vancouver (English) testing rooms were somewhat different. For example, the testing room in Nijmegen was larger than that in Vancouver, though the Nijmegen test booth was approximately the same size as the Vancouver testing room. Also, the experimenter remained in the testing room with the Dutch children (though behind the testing booth, operating the computer, and hence not visible to the participants).

4.8.1 Participants

Our subjects were 35 Dutch children ranging from 17 months 21 days to 19 months, 26 days (mean age: 18 months, 25 days). Participating subjects were again assigned to two groups: Vowel Quality (n=18) and Vowel Length (n=17). 28 additional infants were tested but not included in the final sample because of crying, fussiness, failure to habituate, parental interference during testing, experimenter error. Recruitment took place as described in section 2.3.1.

All children had normal birth records, no history of health problems and exposure to Dutch was at least 70%.

4.8.2 Materials

The stimuli consisted of 20 isolated CVC-word forms of four types, 'tam', 'tem', 'taam and 'niem'; there were 5 tokens of each type. 'taam' tokens were identical to 'tam' tokens, except that the vowel had been lengthened (i.e. doubled in length) using the sound editing program PRAAT, as previously described in section 7.4. The orthographic change from 'neem' in Experiment 4 to 'niem' in Experiment 5 was made for recording purposes only, to keep the pronunciation of our Dutch speaker as similar as possible to the pre-and post-test word form 'neem' in the Canadian experiment. The speaker was a female native speaker of Dutch.

Average measurement for 'tam', 'tem', 'taam' and 'niem' tokens used for the present study are displayed in figure Figure 4.1.

Dutch Stimuli	Mean word duration	Mean vowel length	Mean F1	Mean F2
tam	414.06	154.92	974.50	1318.84
tem	420.156	143.24	844.50	2073.69
taam	556.83	248.39	974.50	1318.84
niem	505.97	137.72	406.93	2721.25

Figure 4.5: Acoustic measurements for Dutch stimuli

The durational differences between the two sets of stimulus materials in Experiment 4 and 5 are discussed in section 4.10.

4.8.3 Procedure

The procedure was the same as Experiment 4 (see section 4.3.3 for details). Just as in the previous experiment, we were interested in whether children would link the two vocalically different novel word forms with the two objects.

4.8.4 Predictions

We predicted that Dutch children would notice the switch in both conditions (see Table 4.2). This prediction was based on 18-month-old children having had considerable listening experience with the phonemically long and short Dutch vowel sounds. Our predictions for Dutch and Canadian children thus differed only for the vowel length condition since this is where the segmental phonology of the two languages differs.

Predictions for Dutch Switch Study (native vowel quality versus vowel duration contrast)			
Vowel Quality	Looking time 'switch' trials	>	Looking-time 'same' trials
Vowel Length	Looking time 'switch' trials	>	Looking-time 'same' trials
• Categorization in line with Canadian segmental phonology			

Table 4.6

4.9 Results

We were primarily interested in whether Dutch children would dishabituate to a switch of minimal pairs differing either in vowel quality or in vowel length.

Figure 4.7 presents the looking time across conditions for children in the Vowel Quality Group and Figure 4.8 for children in the Vowel Length Group. As we had predicted, children in the Vowel Quality Condition looked much longer during the 'switch' as opposed to the 'same' trial. Contrary to the Canadian pattern, a difference in looking time was also apparent for our Dutch children in the Vowel Length Condition.

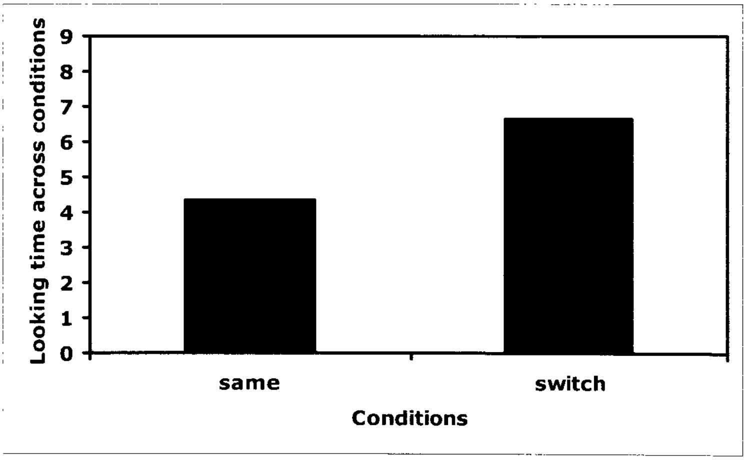

Figure 4.7: Mean looking times across conditions in the Dutch Vowel Quality Switch Study

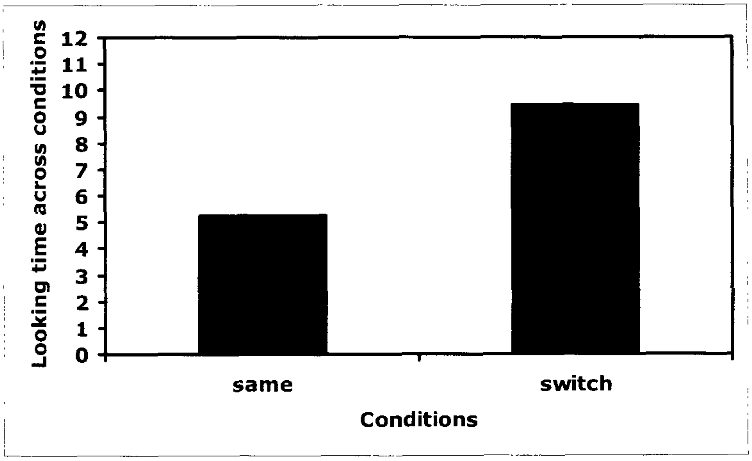

Figure 4.8: **Mean looking times across conditions in the Dutch Vowel Length Switch Study**

As for Experiment 4, it was first determined that children across the two groups maintained their attention throughout the experiment (see section 4.5 for details).

For the Vowel Quality Group, these analyses revealed significant differences between average looking times during the last two habituation trials versus the post-test ($t(17)= 16.99$, $p<0.01$, all tests two-tailed), while no difference was present between the pre-and the post-test ($t(17)= .10$, n. s.).

The same pattern of results emerged for the Vowel Length Group and hence there was no difference between the pre-and the post-test ($t(16)= .79$, n.s.).

This recovery of looking time to pre-test levels during the post-test across the two groups confirms that the Dutch infants were still paying attention during the latter part of the study in both conditions.

Main analyses again consisted of Paired Samples t-tests, comparing average looking times during 'same' and 'switch' trials. For the Vowel Quality Group, this analysis yielded significant differences between 'same' and 'switch' trials ($t(17)= 3.89$, $p= .02$). The same pattern of results was observed for the Vowel Length Group ($t(16)= 4.30$, $p= .01$).

Summing up, Dutch children dishabituated both to the switch of 'tam' versus 'tem' (Vowel Quality Condition) and to the switch of 'tam' versus 'taam' (Vowel Length Condition). As Canadian children noticed only the change in vowel quality but not the change in vowel length, the data from the two populations differ in this regard.

To explore whether there was indeed a difference between the two populations in terms of subjects' processing of vowel length when the task was lexical in nature, the Switch data from the two populations were combined in an (unbalanced) Type I 2-factor ANOVA. Specifically, the differential effect of Language (English versus Dutch) and Contrast (vowel length versus vowel quality) on difference scores was explored.

This analysis revealed a main effect of Language (with greater differences in the Dutch children; $F(1,64)=5.33$, $p<.025$). The interaction between Language and Contrast was significant ($F(1,64)=5.37$, $p<.025$), indicating subjects' native language had a differential effect on their ability to detect the two contrasts.

4.10 Discussion

As noted, Dutch children's looking time pattern differed from that of Canadian children: whereas Canadian children noticed only the switch in the word forms differing in vowel quality Dutch subjects seemed to notice a switch of both vowel quality and vowel length. This pattern of results may be explained in terms of the segmental phonology of the two languages. In Dutch, vowel length as well as vowel quality are phonologically relevant for determining vowel identity. However, in Canadian English, spectral vowel information (i.e. vowel quality) appears to serve as a primary cue to vowel identity, while vowel length may serve as a secondary cue.

As discussed previously, the Vowel Quality Condition was included for control purposes. We were primarily interested in comparing Dutch and Canadian children's processing of the Vowel Length contrast since the role of the latter differs across the two languages. The Vowel Quality Condition served to make sure that there would be other contrasts, native to both populations, that both groups could be expected to process in a phonological way. The latter is indeed what was found in that both Dutch and Canadian children dishabituated to the switch of 'tam' versus 'tem' (Vowel Quality Condition).

It is important to note that there was a durational difference between Canadian English 'tam' and 'tem' tokens both in terms of vowel length and in terms of word duration (i.e. /æ/ was longer than /ɛ/ and hence 'tam' was longer than 'tem') (see Figure 4.1). This difference was not present for the Dutch Vowel Quality contrast (see Figure 4.5). Despite these accompanying durational differences, it is likely that Canadian English children used spectral information as a primary cue when differentiating the two vowels (Hillenbrand et al., 2000) and hence the two word forms, although they may still have drawn upon the durational differences as a secondary cue.

The presence of the durational differences between the two spectrally different vowels would clearly have been problematic if Canadian English children but not Dutch participants had shown positive evidence of processing the word forms 'tam' and 'tem'. However, the latter is not what was found. Hence Dutch children do not appear to have been at a disadvantage compared with Canadian subjects because of the absence of accompanying durational cues in the Vowel Quality Condition.

To conclude, the present data are compatible with the hypothesis that Canadian children processed the two word forms differing in vowel length ('tam' versus 'taam') in a non-phonemic way, unlike Dutch children who showed positive evidence of processing these two word forms as distinct lexical items. In comparison, both groups of subjects showed positive evidence of lexical processing during the Control Condition (i.e. the Vowel Quality Condition).

In comparison with Experiments 2 and 3, which indicated that both Dutch and Canadian infants were able to 'discriminate' word forms differing in vowel length, the present results were obtained under task conditions requiring lexical processing and hence higher level processing. The latter may explain the different pattern of performance in Experiment 4 and 5, in particular Canadian English children's failure to dishabituate to the switch during the Vowel Length Condition in Experiment 4.

5 ACROSS- AND WITHIN-LANGUAGE INFLUENCES ON DUTCH CHILDREN'S ATTENTION TO VOWEL LENGTH

Experiment 6 and 7 were both conducted in the Netherlands with Dutch children to follow up on two questions: First, whether Dutch children would dishabituate when presented with the Canadian stimuli differing in vowel length (Experiment 6), and second, whether Dutch children would dishabituate on a long-short vowel pair for which there is no length opposition in Dutch (Experiment 7).

Experiment 6 was motivated by differences between the stimulus sets across the two languages. Specifically, the short Dutch /ɑ/ tends to be much shorter than the average Canadian /æ/; see also section 4.10). Hence there appear to be intrinsic vowel length differences between Canadian English and Dutch vowels.

An important consideration regarding our stimuli related to the length issue in the preceding paragraph is that the short and long Dutch vowels extended over a different 'length range' from the Canadian stimuli. Hence, although the relative duration difference was identical (i.e. the duration ratio) the absolute length of the short and long vowels across the two language groups was not.

This discrepancy in length may have resulted in at least one potential confound: it is possible that it is harder in general to distinguish and categorize vowels at the long end of the continuum than at the short end of the continuum, irrespective of whether the relative length difference is identical (i.e. the length ratio) (Weber's Law). According to Weber's Law, "for two stimuli of the same general kind to be discriminated the difference between them has to be a constant of the smaller" (Jones, 1974, p.2). If so, Canadian children may have failed to dishabituate to the vowel length variants in Experiment 4, not because of their native language phonology, but because the stimuli were more difficult to tell apart acoustically than the Dutch stimuli. That is, perhaps the Canadian English subjects collapsed the relatively longer Canadian stimuli contrasting in vowel length and categorized them both as 'long', thus not succeeding in associating them with the two animated displays of objects in the Switch task. The relatively shorter Dutch stimuli, however, may have been perceived as short and long by Dutch children. This could explain Dutch participants' success in the word-picture association experiment. If this were the case, the difference in effect we found between the two language groups in the Vowel Length Condition may be, at least in part, an artefact of this perceptual confound.

Importantly though, for the Dutch group, the short-long, as opposed to the long-extra-long contrast was obviously also more familiar based on the ambient language.

If it were indeed harder in general to distinguish and categorize vowels at the long end of the continuum, Dutch children may also have difficulty acquiring two word forms that have the length characteristics of the Canadian stimuli. At the same time, Dutch children's potentially generally heightened attention to vowel length as a result of their native language experience may allow them to compensate for the 'length confound' to some extent, in which case they may still succeed in such a task.

Experiment 6 served to address this question by testing Dutch children's ability to acquire the Canadian 'tam'-'taam' word forms, using the Switch Paradigm.

Another consideration pertaining to the Dutch group is that Dutch participants' dishabituation during the Vowel Length Condition in Experiment 5 is not proof that Dutch children will be able to 'acquire' any native minimal pair differing only in vowel length in the Switch procedure. In other words, children's success may require the vowels to participate in the length opposition (see section 3.1 for a discussion of which vowels do and which do not have long-short counterparts in Dutch).

Experiment 7 served to address this question by testing Dutch children's ability to acquire two word forms differing in vowel length, where the short variant word was a native short vowel of Dutch that does not have a long counterpart. Hence the word form containing the long variant of this vowel should not qualify as a distinct lexical item for Dutch children.

If 17 to 18-month-old Dutch children only processed vowel length contrastively when presented with Dutch vowels for which there is a short-long contrast in Dutch they would be expected not to dishabituate in Experiment 7. Positive evidence of dishabituation in Experiment 7, however, would suggest that Dutch children process vowel length contrastively, irrespective of whether there is a short-long contrast in Dutch.

5.1 Experiment 6

Experiment 6 served to explore the question of whether the Canadian English stimuli in the Vowel Length Condition may have been more difficult to acquire because of their longer relative durations compared with the Dutch stimuli. To address this question a replication of the Dutch Switch study was undertaken, but with the Canadian stimulus materials (i.e. the Canadian 'tam'-'taam' word forms) from Experiment 4.

5.2 Method

The Switch Paradigm was used for the present experiment (for procedural details, see section 4.3.3). The experiment was almost identical to Experiment 4; the only difference – apart from a new group of subjects being tested – was that the stimuli were the Canadian stimuli used in Experiment 4.

5.2.1 Participants

Our subjects were 17 Dutch children ranging from 17;0 months to 18;0 months (mean age: 17;10 months). Eight additional infants were tested but not included in the final sample because of crying, fussiness, failure to habituate, parental interference during testing, equipment- or experimenter error. Recruitment took place as described in section 2.3.1.

All children had normal birth records, no history of health problems and exposure to Dutch was at least 70%.

5.2.2 Materials

The stimuli were Canadian English word forms, identical to those described in section 4.3.2. We tested Dutch participants on the vowel length contrast only and not on the quality contrast since it was in the former condition that we had found different effects between the two language groups (i.e. Experiment 4 versus 5). The stimuli, therefore, consisted of 15 rather than 20 isolated CVC-word forms. The word form types were 'tam', 'taam' and 'neem'; there were 5 tokens of each type.

5.2.3 Procedure

The procedure was the same as in Experiment 4 and 5 and will therefore not be described again (see section 4.3.3 for details). In the present experiment, we were interested in whether Dutch children would link the two Canadian word forms differing only in vowel length with the two animated displays of novel objects in the Switch task.

5.2.4 Predictions

We predicted that Dutch children would still notice the switch of the word form containing the short vowel variant to its long vowel counterpart (see Figure 5.1). The rationale for this prediction was that we were still dealing with a contrast in vowel length to which we predicted Dutch children to have heightened attention. However, failure to dishabituate was also conceivable; after all, the word forms now contained non-native vowels, although the shorter /æ/ may have been assimilated to the Dutch /ɛ/. This tendency to collapse English /æ/ to Dutch /ɛ/ has been observed for Dutch adult listeners when listening to British English /æ/ (Broersma, 2005a; also see Weber & Cutler, 2004).

Predictions for Dutch Switch Study (Canadian stimuli)			
Vowel Length	Looking time 'switch' trials	>	Looking time 'same' trials
• Categorization in line with assumption that Dutch infants' perception of vowel length is generally heightened			

Figure 5.1

5.3 Results

We were interested in whether Dutch children would dishabituate to a switch of minimal pairs differing exclusively in vowel length when presented with the Canadian stimuli used in Experiment 5. Figure 5.2 presents the looking times across conditions for infants in the present study.

Children in the present experiment dishabituated during the 'switch' trial, even though Canadian children had not done so when listening to the very same stimuli.

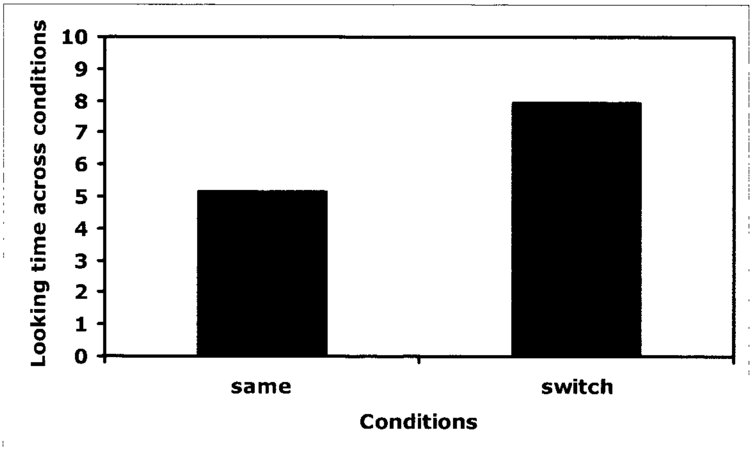

Figure 5.2: **Mean looking times across conditions in the Dutch Vowel Length Switch Study with Canadian stimuli**

Just as for Experiment 4 and 5, it was first determined that children across the two groups maintained their attention throughout the experiment (see section 4.5 for details).

For the children in the present experiment, these analyses revealed significant differences between average looking times during the last two habituation trials versus the post-test ($t(16)= 23.35$, $p<0.01$, all tests two-tailed), while no difference was present between the pre-and the post-test ($t(16)= .07$, n. s.).

This recovery of looking time to pretest levels during the post-test confirms that the Dutch children were still paying attention during the latter part of the study.

Main analyses consisted of Paired Samples t-tests, comparing average looking times during 'same' and 'switch' trials. This analysis yielded significant differences between 'same' and 'switch' trials ($t(16)= 3.58$, $p<.03$).

Summing up, Dutch children in this experiment dishabituated to the switch of vowel length in the novel word forms when listening to Canadian stimulus words.

5.4 Discussion

Dutch children's looking time pattern in the present experiment suggests that they succeeded in associating the two novel word forms differing in vowel length with their associated animated displays of novel objects. Whereas Canadian children failed to notice the switch of the word forms differing in vowel duration in Experiment 4, Dutch children did not, even though they were tested on the exact same stimuli as the Canadians.

The present data thus suggest that the relative length of the vowels in Experiment 4 and 5 was not the reason for Canadian children's failure to dishabituate to the 'switch' trial in the Vowel Length Condition in Experiment 4.

Moreover, the present data provide strong corroborating evidence for the hypothesis that Dutch children may be developing perceptual capacities that allow them to categorize word forms differing in vowel length as distinct word forms, irrespective of whether the target vowels extend over the short or the longer range of the vowel length continuum.

Moreover, the data suggest that Dutch children may process vowel length lexically even in cases where the target vowel types are non-native. Might it be the case that children exposed to a language with a phonemic vowel length difference such as Dutch are developing perceptual sensitivity to the vowel length parameter 'in general' by about 18 months? This could explain why Dutch children dishabituated both when listening to Dutch and when listening to Canadian stimuli.

Alternatively, Dutch children may have found it easier than Canadian children to process the vowel length difference between the Canadian stimulus words because the stimuli were non-native in nature; just as Japanese L2-speakers of English seem to pay attention to vowel length when listening to non-native English vowels, native Dutch listeners might also attend to vowel length relatively more than they would if the vowel length difference were a native Dutch vowel that does not participate in the vowel length opposition.

Experiment 7 served as a first attempt to address this question.

5.5 Experiment 7

Having established that Dutch infants show evidence of categorizing two word forms containing phonemically long and short Dutch vowels (i.e. Experiment 2) as well as words containing non-native long and short vowels that extend over the longer range of the vowel length continuum (Experiment 5), it still remained unclear whether Dutch children's lexical processing of vowel length variation may be restricted to specific vowel contrasts. For example, it may be that this perceptual sensitivity is limited to the aforementioned types of vowel contrasts, namely those, for which there is a length distinction in Dutch and those that are non-native in nature. As noted, there are vowel types in Dutch for which there is no clear length variant. The short Dutch vowel /ɛ/ belongs to this group of vowels and hence it has no long counterpart.

If Dutch children had indeed developed general perceptual sensitivity to vowel length, their sensitivity should also extend to vowels for which there is no long-short distinction in Dutch such as the aforementioned short Dutch vowel /ɛ/.

Thus we tested whether Dutch children would succeed in learning the two novel word forms /tɛm/ and /tɛ:m/, even though /tɛ:m/ should not qualify as lexically different from /tɛm/ for Dutch children since the long vowel /ɛ:/ does not exist in Dutch.

5.6 Method

The Switch Paradigm was employed for the present experiment in the same way as in Experiments 4 to 6. Details of the procedure may be found in section 4.3.3.

5.6.1 Participants

Our subjects were 16 Dutch children ranging from 17; 8 months to 18; 14 months (mean age: 17 months, 18 days). Nine additional infants were tested but not included in the final sample because of crying, fussiness, failure to habituate, parental interference during testing, experimenter error. Recruitment took place as described in section 2.3.1.

All children had normal birth records, no history of health problems and exposure to Dutch was at least 70 %.

5.6.2 Materials

There were two target stimulus word forms as well as the previously used pre-and post-test word 'niem'. The target word forms were the Dutch word 'tem' (the quality control word from Experiment 4) and its lengthened variant 'teem'. The stimulus word 'teem' was created in the same way as the lengthened word forms in the previous experiments, namely by doubling the length of the vowel of its short counterpart (i.e. 'tem'), using PRAAT-software (see section 7.4). We used the already existing Dutch recording of the words 'tem' and 'niem' and hence no change of speaker was necessary. The only new stimulus was hence the lengthened 'teem'.

5.6.3 Procedure

We used the Switch procedure in the same way as described in section 4.3.3.

5.6.4 Predictions

As indicated in Figure 5.3, we predicted that Dutch children would still notice the switch from the word form containing the short vowel variant /ɛ/ to its long counterpart /ɛ:/ since we were still dealing with a contrast in vowel length. However, failure to dishabituate was also considered possible, considering that there is no long-short opposition for the vowels under investigation in the present experiment in the infants' native segmental phonology (i.e. Dutch).

Predictions for Dutch Switch Study (/tɛm/ versus /tɛ:m/)			
Vowel Length	Looking time 'switch trials'	>	Looking time 'same trials'
• Categorization in line with assumption that Dutch children may be developing a general vowel length parameter independent of vowel type/category			

Figure 5.3

5.7 Results

We were primarily interested in whether Dutch children would dishabituate to a switch of minimal pairs differing exclusively in vowel length where the stimuli contained Dutch vowels for which only the short vowel exists in the ambient language (i.e. the short vowel has no long counterpart in Dutch).

Figure 5.4 presents subjects' looking times across conditions. As we had predicted, Dutch children's looking time increased during the 'switch' relative to the 'same' trial.

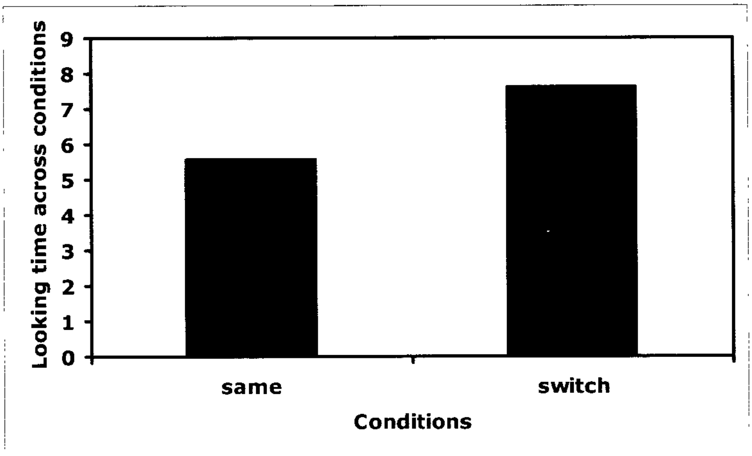

Figure 5.4: **Dutch Switch Study with Native - Non-native Dutch Vowel Length Contrast**

Again, it was first determined that children maintained their attention throughout the experiment (see section 4.5 for details).

For the children in the present experiment, these analyses revealed significant differences between average looking times during the last two habituation trials versus the post-test ($t(15)= 14.10$, $p<0.01$, all tests two-tailed), while no difference was present between the pre-and the post-test ($t(15)= .26$, n. s.).

This recovery of looking time to pretest levels during the post-test confirms that the Dutch children were still paying attention during the latter part of the study.

Main analyses again consisted of Paired Samples t-tests, comparing average looking times during 'same' and 'switch' trials. This analysis yielded significant differences between 'same' and 'switch' trials ($t(15)= 2.15$, $p<.05$).

Summing up Dutch children dishabituated to the switch of vowel length when tested on the native word form /tɛm/ and its non-native length variant /tɛ:m/.

5.8 Discussion

Dutch children's looking times across conditions in Experiment 7 were similar to Dutch subjects' pattern of performance in Experiment 4. To recapitulate, Experiment 4 employed a an existing Dutch vowel length contrast between /ɑ/ versus /a:/. Both groups of children hence noticed the mapping violation and increased their looking times during the 'switch' trial.

The fact that Dutch children across the two experiments noticed the switch indicates that they were paying attention to vowel length in general, as opposed to merely attending to it when tested on a short Dutch vowel for which there is a contrasting long vowel in the Dutch language (Experiment 4).

The hypothesis that Dutch children may have developed a general vowel length parameter is corroborated by our findings obtained in Experiment 6. In the latter, Dutch children were found to categorize the non-native Canadian length variant words 'tam' and 'taam' as two distinct words, unlike Canadian children in Experiment 4.

Dutch children's attention to phonetic vowel length variation when processing non-native stimuli (as in Experiment 6) may be due to their having learned that vowel length can signal a phonemic distinction for vowels in the ambient language. As they do not know what role vowel length plays in Canadian English since they have not been exposed to Canadian English, paying attention to vowel length when listening to Canadian English stimuli may thus reflect the use of a native listening strategy when listening to non-native stimuli.

When compared with Experiment 4 (Canadian Switch Study), the present data with Dutch children provide support for the hypothesis that Dutch children's perceptual processing may be in the process of developing into language-specific sensitivity only insofar as Dutch children pay relatively more attention to vowel length contrasts between otherwise identical word forms (i.e. minimal pairs) than Canadian children. At the same time, Dutch children do not appear to treat vowels that do and that do not participate in the length opposition in the ambient language differently. If they did, we would not have expected them to have dishabituated when tested on 'tem' versus 'teem' in Experiment 7 because these two word forms should not qualify as two distinct lexical items according to the phonology of Dutch.

To conclude, evidence has been obtained suggesting that Dutch 17-month-old children treat vowel length contrasts lexically. The latter appears to be applicable even when the children hear vowels for which there is no short-long opposition in the ambient language and when they hear non-native stimuli.

6 SUMMARY AND CONCLUSIONS

6.1 Distinguishing Contrastive from Non-Contrastive Variation

This dissertation explored infants' learning about phonologically relevant and irrelevant phonetic variation in speech perception. The test case used here was vowel length. Specifically, we compared the processing of vowel length variation cross-linguistically in infants learning Dutch and infants learning Canadian English.

Vowel length is phonemic in Dutch but not in Canadian English. Hence Dutch infants should eventually classify two minimal pair word forms that differ only in vowel duration (e.g. 'tak' versus 'taak') as two distinct words rather than two realizations of the same word. Canadian English infants, however, should eventually classify two minimal pair words such as 'cat' and 'caat' as two realizations of the same word, although they may still perceive the word form 'caat' as an unusual realization of the more prototypical realization 'cat'. They may thus still be able to discriminate 'cat' and 'caat'.

When infants learn about the role that vowel length variation plays in their native language, they face a major challenge: Vowel length is affected by several contextual factors, including speaking rate, stress, the word's position in the sentence and the surrounding phonetic context. For example, in English, vowel length serves as a critical cue to determining the voicing status of the following plosive such that vowels preceding voiced plosives are longer compared with vowels preceding voiceless plosives (e.g. 'tab' versus 'tap'). Consequently, English-learning infants cannot simply ignore vowel length variation altogether since they would otherwise miss an important source of information to postvocalic consonant voicing. Yet, English infants do have to learn that vowel length does not serve as a primary cue to vowel identity and that they should pay relatively more attention to vowel quality as opposed to durational differences between vowels. In Dutch, vowel length is phonemic only for some vowels (e.g. /ɑ/ versus /a:/) but not for others (e.g. /ɛ/ versus /ɛ:/). Thus Dutch infants have to learn to process vowel length phonemically for those vowels that participate in the length opposition and ignore variation in length for those that do not.

In this dissertation I investigated whether young children have learnt about the role that vowel length plays in the ambient language at one level of analysis – that is, whether it serves as a primary cue to vowel identity. Our results suggest that 17-month-olds have indeed learnt about the phonemic status of vowel length in the ambient language (see the experimental summary below).

6.2 Resilience of Early Lexical Representations to Non-Contrastive Acoustic Variation

While it has been demonstrated in previous research that infants' phonemic representations at the individual sound and syllable level are language-specific towards the end of the first year of life, most studies examining language-specificity at the word level have used phonemic rather than non-phonemic sound contrasts. Here, we used a contrast that was phonemic for one group of infants and non-phonemic for the other.

By using a cross-linguistic approach and comparing two language groups, for which vowel length was either phonemic or non-phonemic, we were able to make different predictions about infants' processing of vowel length variation based on the segmental phonology of the two languages. Therefore, differential processing of vowel length variation by Dutch and Canadian infants in line with subjects' native segmental phonology would suggest that infants' early lexical representations may be well-specified vocalically and would further suggest that these representations contain language-specific detail about durational aspects of vowels. This is exactly what we found (see 6.4).

6.3 Dealing with Salient but Non-Phonemic Acoustic Variation

Previous studies that have explored the processing of non-native vowel distinctions in infants have mostly investigated more subtle differences between vowels. In the present set of experiments, we focused on a different question: We asked how infants process salient acoustic variation of vowel duration that is in the signal but nonetheless non-phonemic for Canadian infants.

Both adults and children and even newborns are known to be sensitive to very small durational differences of stimuli for different aspects of speech processing (e.g. speaking rate compensation). The fact that even newborns show sensitivity to this temporally-based acoustic variation suggests that sensitivity to this source of variation may be independent of language experience, at least in part. Alternatively, the acoustic salience of vowel length differences could be another reason for listeners' sensitivity to this variation. Finally, listeners' ability to perceive small durational variations of stimuli may also be influenced by the importance of vowel length variation at several levels of processing across many languages.

Presumably, how listeners interpret variations in the length of vowels is ultimately shaped by the ambient language, as we have shown in this thesis (see 6.4).

6.4 Experimental Findings

Experiment 1 explored the specificity of infants' early lexical representations using the conditioned headturn method (CHT). Dutch-learning eight to twelve-month-old infants were trained to make a headturn response to a novel word form. During testing, the trained target and vocalically varying foils were embedded in sentence contexts and infants' headturn behaviour was monitored. Foils varied either in vowel quality or in vowel length. The vowel length difference was phonemic, the vowel quality difference was not. Infants' headturn responses to the target as opposed to the similar sounding foils was taken as evidence of increased activation of the trained target relative to its variants. Although proportional differences in headturns across conditions did not differ significantly in the present experiment, response latencies did. Specifically, Dutch infants turned fastest overall to the trained target and significantly more slowly to the phonemically different length variant but not to the non-native vowel quality variant. One way to interpret this finding is to attribute the differences in response latencies to Dutch segmental phonology: If Dutch infants' responses were shaped by the ambient language they would be expected to perceive the trained target 'hasp' as lexically different from its length variant word 'haasp' but not from its quality variant word 'husp'. Another possible reason for the pattern of results obtained is that the non-phonemic vowel quality difference was more subtle acoustically than the native vowel length difference.

To help us determine whether cross-linguistic differences exist in the processing of vowel length we subsequently compared its processing across two populations, one in which vowel length is phonemic (Dutch) versus one in which vowel length is non-phonemic. Before being able to explore the processing of vowel length variation at the lexical level, it was first important to determine that infants from both populations would be able to discriminate stimuli varying in vowel duration.

Therefore, we employed a discrimination task in Experiments 2 and 3 and compared seven-and-a-half to twelve-month-old Canadian English-and Dutch-learning infants' differential ability to recognize trained word forms in utterance context, again using the CHT-method. Variant words differed either in vowel quality (phonemic for both populations) or in vowel length (phonemic only for Dutch infants). Given the different role vowel length plays across the two languages, we predicted that Dutch infants would turn relatively less both to the vowel quality and to the vowel length variant. In contrast, Canadian infants were predicted to turn relatively less to the phonemic vowel quality difference but not to the non-phonemic vowel length variant word. Our findings suggested otherwise: Both groups of infants turned significantly less to both variant words, thus suggesting that both Dutch and Canadian English infants had categorized both the quality variant and the vowel length variant as different word forms from the trained target.

The Canadian result may seem surprising given the non-phonemic status of vowel length in Canadian English. However, a possible explanation for the above findings may be the nonreferential nature of the CHT methodology. The results may also have been influenced by the acoustic salience of vowel length. For infants to succeed in the CHT-task, they had to form a detailed representation of the trained target word, retain this representation for the duration of the test phase and subsequently make acoustic similarity judgements of the stimuli. No word-object associations were necessary, however. Consequently, infants did not need to draw upon a lexical level of processing. Instead, a lower level of processing at the acoustic level may have sufficed. The latter may explain why Canadian infants were sensitive to vowel length variation and discriminated the length variant words despite the non-phonemic status of vowel length in the ambient language. It is also possible that (salience aside) infants responded on the basis of the same perceptual sensitivities that are necessary for speech interpretation at other linguistic levels. Vowel length is a factor in, for example, perceiving English (and Dutch) word stress, identifying sentence accent, and of course categorizing voicing in syllable-final consonants.

To conclude, we demonstrated that Dutch and Canadian English infants are able to recognize a trained word form in utterances and that they distinguish this word form from variant words differing in vowel quality and vowel length. Critically, even Canadian infants, for whom vowel length is non-phonemic, demonstrated this ability. Considering the importance of vowel length information at other levels of processing in English, however, being sensitive to variation in vowel duration appears not only useful but critical for Canadian infants as a prerequisite to becoming efficient at compensating for natural variation in everyday speech (e.g. speaking rate normalization).

The findings of Experiments 2 and 3 raised the question of whether language-specific differences in the processing of vowel duration between Dutch and English infants would emerge if older infants were tested using a more referential task. Such a task may induce English infants to call upon a more lexical level of processing, which in turn may lead them to ignore vowel length variation for the purpose of determining vowel identity.

Therefore, Experiments 4 and 5 aimed to test whether the use of a more referential task would induce language-specific differences across the two populations in line with the segmental phonology of the languages. Specifically, we asked whether Dutch and Canadian 17-month-old children would be able to learn to map two word forms differing only in vowel duration onto two different objects. If Dutch but not Canadian English children were able to learn the two words, this would provide strong evidence for 17-month-old children's lexical representations being well-specified for vowel duration in line with the segmental phonology of the native language.

The two groups of children were familiarized with two word-object associations. We subsequently examined whether children would notice a switch in the pairing of the words and the objects.

In contrast to the previous set of studies, a cross-linguistic difference emerged in the present experiments in that Dutch but not Canadian English children noticed the switch in vowel length, whereas both populations noticed a switch in vowel quality (native to both groups), which was our control measure.

To conclude, the present data corroborate the hypothesis that the use of a lexical-level task can induce children to make use of a more language-specific mode of processing in that English infants no longer seem to pay attention – or to notice – the vowel length variation in the more referential Switch task. Even though we did not specifically test English-learning 17-month-olds' ability to discriminate words differing in vowel length (i.e. we tested much younger infants in Experiments 2 and 3), it would be hard to imagine that they would not be sensitive to such variation given the importance of vowel length information at other levels of processing in English and across many other languages and considering the published reports of newborns' and adults' sensitivity to vowel duration manipulations noted above.

Experiments 6 and 7 served to address two crucial questions motivated by the above set of studies: First we explored whether Canadian infants' failure to learn to associate the two novel word forms differing in vowel length onto animated displays of objects in the Switch task may have been due to differences in the stimuli across the two languages. Specifically, the Dutch stimulus words contained much shorter vowels than the Canadian words, although the vowel length ratio between the long and short variant words was the same (2:1). The vowel duration differences were present because they reflected natural variation of the length of the vowels in IDS across the two languages. As it is possible that the relatively longer Canadian stimuli were easier to discriminate, not because of the subjects' native segmental phonology, but because of their acoustics (i.e. it may be that vowel length contrasts are easier to discriminate at the long end of the vowel length continuum).

Therefore, Experiment 6 tested Dutch infants' ability to learn to associate the two Canadian length variants with two novel objects in the Switch task. Dutch infants were successful in learning the two word forms, just as they had been with Dutch stimulus words. This suggests that the natural length differences between the two sets of stimulus materials were not a confound in Experiments 3 and 4.

Experiment 7 was motivated by the question of whether Dutch infants would process vowel length contrastively for all vowels of Dutch or only for those that participate in the length opposition (i.e. not all short Dutch vowels have a long counterpart). We addressed this question by testing Dutch infants' ability to learn the two novel words 'tem' and 'teem'. The /ɛ-ɛː/ distinction is not present in Dutch child-directed speech, though the long /ɛː/ does exist in certain rare loan words. Positive evidence of infants' learning the two words would be suggestive of infants' attention to vowel length variation not being restricted to those vowels of Dutch for which vowel length is phonemic.

Results indicated that Dutch infants were indeed able to learn to associate the two words with the two novel objects despite the vowel length contrast not being part of the Dutch vowel length opposition. This raises the question of whether Dutch infants may be acquiring a general vowel length parameter that induces them to heighten their attention to vowel length in general and not exclusively when processing vowels for which there is a short and long counterpart in Dutch. The fact that Canadian 17-month-olds failed to learn words differing in vowel duration in the Switch task suggests that we are indeed dealing with a language-specific effect. If acoustic salience were critical here, Canadian subjects should also have succeeded at learning to map minimal pairs differing in vowel duration onto novel objects in the Switch task, which they did not.

6.5 Conclusion

This dissertation explored the learning of contrastive and non-contrastive acoustic variation during phonological acquisition: Specifically, we compared the processing of vowel length variation cross-linguistically in two populations, in which vowel length variation is either phonemic or not. Critically, vowel length provides important information at other levels of processing in both language groups. Both discrimination and word-learning tasks were employed such that infants' processing of vowel length variation could be assessed at different levels of processing and under different task demands.

We demonstrated that both Dutch and Canadian English learning infants are sensitive to vowel length differences at eight to twelve months of age and that they are able to draw upon this sensitivity during discrimination tasks.

Crucially, at 17 months, a cross-linguistic difference emerged when the two language groups were tested in a word learning task such that Dutch but not English children made use of vowel length variation during word learning.

To conclude, these findings are in line with the segmental phonology of Dutch and Canadian English. Whereas successful discrimination appears critical for both Dutch and English children at both ages due to the importance of vowel length information for other aspects of processing in both languages, paying attention to vowel duration during word learning is critical only for Dutch but not for English children. Both abilities are reflected in our data. It appears, therefore, that 17-month-old children's knowledge of phonological structure is already well-specified for vowel length, in accordance with the segmental phonology of the respective languages.

7 APPENDICES

7.1 Materials for Experiment 1

Four different tokens of each of the training words were used. Each of the test sentences was recorded with 'hasp', 'haasp' and 'fiem' in utterance-final position.

Isolated words (training phase)	
Background words:	kief, tesp, peets, tiek, fuusp, tief, hets, hierk and pesk
Target word:	hasp
Variant words:	haasp, husp, fiem

Utterances (test phase)
Heeft Lisa trek in een klein hasp/haasp/husp/fiem?
Lust je een ook een lekker tonijn hasp/haasp/husp/fiem?
Heb je zin in een feest met Martijn Hasp/Haasp/Husp/ Fiem?
Hoe vind Susi zijn hasp/haasp/husp/fiem?
Rita bestelt iedere week een fijn hasp/haasp/husp/fiem.
Kirsten en Peter lezen een fijn hasp/haasp/husp/fiem.
Je vindt de doos zeker onder zijn hasp/haasp/husp/fiem.
Leni vindt snel een fijn hasp/haasp/husp/fiem.
Meneer Leuk zoekt plein Hasp/Haasp/Husp/Fiem.
De ober eet het liefst met mijn hasp/haasp/husp/fiem.
Simone eet 's ochtens een fijn hasp/haasp/husp/fiem.
De eend eet liever een klein hasp/haasp/husp/fiem.
Kirsten kent me niet zonder mijn hasp/haasp/husp/fiem.
Peter rent sneller met zijn hasp/haasp/husp/fiem.
De kelner heeft wel een klein hasp/haasp/husp/fiem.

Utterances (test phase)
De oude deur heeft weer een klein hasp/haasp/husp/fiem.
Jullie hebben een fijn hasp/haasp/husp/fiem.
In de tuin vind je een fijn hasp/haasp/husp/fiem.
De wezel lijkt op een tonijn hasp/haasp/husp/fiem.
Mijn lievelingsbloem is een fijn hasp/haasp/husp/fiem.
Ik wil de kok helpen met zijn hasp/haasp/husp/fiem.
Soms win je er een klein hasp/haasp/husp/fiem.
Hij drinkt nooit iets zonder zijn hasp/haasp/husp/fiem.
Je ziet hem zelden zonder zijn hasp/haasp/husp/fiem.
Probeer snel een stuk zwijn hasp/haasp/husp/fiem.
Fiets snel naar plein Hasp/Haasp/Husp/Fiem.
Hij is ook bij het Rijn hasp/haasp/husp/fiem.
De nieuwe pianist is Stijn Hasp/Haasp/Husp/Fiem.
Ze speelt met Hein Hasp/Haasp/Husp/Fiem.
Hij kent Marjolein Hasp/Haasp/Husp/Fiem.
Onze beste docent is Hein Hasp/Haasp/Husp/Fiem.
De witte bloes stuur ik aan Hein Hasp/Haasp/Husp/Fiem.
Ze leert Engels het liefst bij Marjolein Hasp/Haasp/Husp/ Fiem.

7.2 Materials for Experiment 2

Four different tokens of each of the training words were used.

Each of the test sentences was recorded with /tɑk/ and /tɛk/ in utterance-final position. The target word /tɑk/ was then manipulated thus creating 'taak'-sentences, as discussed in section 3.3.3 of this chapter.

Isolated words (training phase)
Background words: tik, tok, ket
Variant words: taak (/tɑ:k/), tek (/tɛk/)
Target word: tak (/tɑk/)

Utterances (test phase)
Rian beschrijft een tak/taak/tek
Peter tekent een tak/taak/tek
Meneer Broes zoekt een tak/taak/tek
Simone vindt een tak/taak/tek
In de tuin vind je een tak/taak/tek
De kelner toont hem een tak/taak/tek
De wezel eet nooit een tak/taak/tek
Het kind speelt met een tak/taak/tek
Iedere boom heeft een tak/taak/tek

7.3 Materials for Experiment 3

Isolated words (training phase)
Background words: tick, tock, cet
Variant words: taack (/tæ:k/), teck (/tɛk/)
Target word: tack (/tæk/)

Utterances (test phase)
Megan showed me the tack/taack/teck
Mr. Shi prefers the tack/taack/teck
The waiter pours the tack/taack/teck
The teacher knows the tack/taack/teck
Ella slid across the tack/taack/teck
Jeremy prefers the tack/taack/teck
Linda skates on the tack/taack/teck
The pianist performs the tack/taack/teck
The finch enjoyed the tack/taack/teck

7.4 Manipulation of Duration in PRAAT

Open PRAAT

Click on 'Read' followed by 'Read from file..' and open sound file to be manipulated (e.g. word form 'tak') in wave-format.

Click on the sound-file 'Praat objects window'.

Click on 'To Manipulation'... followed by 'OK'. This opens 'Manipulation object' in 'Praatobjects' window. Click on 'Manipulation object' (e.g. Manipulation 'tak').

Then click on 'Edit'; this opens waveform, pitch tier and duration tier.

Zoom into vowel and move curser to beginning of vowel at 'zero-crossing point'.

Click on 'Dur' followed by 'Add duration point at..'.

Pick 'duration manipulation factor' (e.g. 2 to double duration); this needs to be done at offset of vowel as well for it to apply to vocalic portion only rather than remainder of sound-segment (i.e. consonant).

8 REFERENCES

Abrams, R. M., Gerhardt, K. J., Huang, X., Peters, A. J. M., & Langford, R. G. (2000). Musical experiences of the unborn baby. *Journal of Sound and Vibration*, 231, 253-258.

Akamatsu, T. (1997). *Japanese Phonetics: Theory and Practice.* Muenchen & Newcastle: Lincom Europa.

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, 38, 419-439.

Assmann, P. F. T., Nearey, T. M., & Hogan, J. (1982). Vowel identification: Orthographic, perceptual and acoustic aspects. *Journal of the Acoustical Society of America*, 71, 975-989.

Aslin, D., Jusczyk, P.W., & Pisoni, D. B. (1998). In: D. Kuhn & R. S. Siegler (Eds.), *Handbook of Child Psychology*. 5th ed., vol. 2: *Cognition, perception and language*. New York: John Wiley & Sons, Inc.

Avis, W. S. (1975). The phonemic segments of an Edmonton idiolect. In: J. K. Chambers (Ed.). *Canadian English: Origins and Structures* (pp. 118-128). Toronto: Methuen.

Barton, D. (1976). Phonemic discrimination and the knowledge of words. *Papers and Reports on Child Language Development*, 11, 61-68.

Barton, D. (1978). The discrimination of minimally- different pairs of real words by children aged 2;3 to 2;11. In: N. Waterson, & C. Snow (Eds.), *The Development of Communication* (pp. 255-261). Chichester: Wiley.

Best, C. T. (1994). Learning to perceive the sound pattern of English. In: C. Rovee-Collier, & L. Lipsitt (Eds.), *Advances in Infancy Research* (pp. 217-304). Hillsdale, NJ: Ablex.

Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganisation for non-native speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 345-360.

Bohn, O.-S. (1995). Cross-language speech perception in adults: First language transfer doesn't tell it all. In: W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* (pp. 279-304). Baltimore: York Press.

Brent, M. R., & Siskind, J. M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition*. 81. B33-B44.

Boersma, P., & Weenink, D. (2005). Praat: Doing Phonetics by Computer [Computer Program]. Retrieved from http://www.praat.org/.

Botinis, A., Banner, R., Fourakis, M., & Pagoni-Tetlow, S. (2002). Prosodic effects and crosslinguistic segmental durations. *Fonetik*, 44, 77-80.

Broersma, M. (2005a). Phonetic and lexical processing in a second language. Unpublished doctoral dissertation, University of Nijmegen (MPI Series in Psycholinguisitics, 34).

Broersma, M. (2005b). Perception of familiar contrast in unfamiliar positions. *Journal of the Acoustical Society of America*, 117, 3890-3901.

Buder, E. H., & Stoel-Gammon, C. (2002). American and Swedish children's acquisition of vowel duration: Effects of vowel identity and final stop voicing. *Journal of the Acoustical Society of America*, 111, 1854-1864.

Cohen, L. B. (1992). The myth of differentiation. Symposium paper presented at the International Conferences on Infant Studies. Providence, RI.

Christophe, A., Dupoux, E., Bertoncini, J., & Mehler, J. (1994). Do infants perceive word boundaries? An experimental study of the bootstrapping of lexical acquisition. *Journal of the Acoustical Society of America*, 95, 1570-1580.

Christophe, A., & Morton, J. (1998). Is Dutch native English? Linguistic analysis by 2-month-olds. *Developmental Science*, 1, 215-219.

Cross, T. G. (1977). Mother's speech adjustment : the contributions of selected child listener variables. In: C. E. Snow, & C. A. Fergusson (Eds.), *Talking to Children: Language Input and Acquisition* (pp. 155-188). Cambridge: Cambridge University Press.

Crosswhite, K., & Tanenhaus, M. K. (2004). Online use of phonetic detail in spoken word recognition. *9th Conference on Laboratory Phonology*, June 24–26.

Crystal T. H., & House, A. S. (1988). The durations of. American English vowels: An overview. *Journal of Phonetics*, 16, 263-284.

Cutler, A. (1990). Exploiting prosodic probabilities in speech segmentation. In: G. Altmann (Ed.), Cognitive Models of Speech Processing: *Psycholinguistic and Computational Perspectives* (pp. 105-121). Cambridge (MA): MIT Press.

Cutler, A. (2000). Listening to a second language through the ears of a first. *Interpreting*, 5 (2000/2001), 1-23.

Cutler A., & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language*, 2, 133-142.

Cutler, A., & Mehler, J. (1993). The periodicity bias. *Journal of Phonetics*, 21, 103-108.

Cutler, A., & Norris, D.G. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 113-121.

DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: newborns prefer their mother's voices. *Science*, 208, 1174-1176.

DeCasper, A. J., & Spence, M. J. (1986). Prenatal maternal speech influence on newborns' perception of sounds. *Infant Behaviour and Development*. 9, 133–150.

Eimas, P. D., Siqueland, E. R., Jusczyk, P.W., & Vigorito, J. (1971). Speech perception in infants. *Science*, 171, 303-306.

Erickson, M. L. (2000). Simultaneous effects on vowel duration in American English: A covariance structure modeling approach. *Journal of the Acoustical Society of America*, 108, 2980-2995.

Fennell, C. T., & Werker, J. F. (2003a, October). Applying speech perception to word learning at 14 months: The effects of word knowledge and familiarity. *Boston University Conference on Language Development*, Boston, USA.

Fennell, C. T., & Werker, J. F. (2003b). Early word learners' ability to access phonetic detail in well-known words. *Language and Speech*, 46, 245-264.

Fernald, A., Swingley, D., & Pinto, J. P. (2001). When half a word is enough: infants can recognize spoken words using partial phonetic information. *Child Development*, 72, 1003-1015.

Frauenfelder, U., & Floccia, C. (1998). The recognition of spoken words. In: A. P. Friederici (Ed.), *Language Comprehension: A Biological Perspective* (pp. 1-40). Berlin: Springer-Verlag.

Gregg, R. J. (1975). The phonology of Canadian English as spoken in the area of Vancouver, British Columbia. In: J. K. Chambers (Ed.), *Canadian English: Origins and Structures* (pp. 118-128). Toronto: Methuen.

Grieser, D. L., & Kuhl, P. (1988). Maternal speech to infants in a tonal language: support for universal prosodic features in motherese. *Developmental Psychology*, 24, 14-20.

Grieser, D., & Kuhl, P. K. (1989). Categorisation of speech by infants: Support for speech-sound prototypes. *Developmental Psychology*, 25, 577-588.

Gussenhoven, C. (1999). Dutch. In: Handbook of the International Phonetic Association. A Guide to the Use of the International Phonetic Alphabet (pp. 74-77). Cambridge: Cambridge University Press.

Hallé, P., & Boysson-Bardies, B. (1986). The format of representation of recognized words in infants' early receptive lexicons. *Infant Behaviour and Development*, 19, 463-481.

Hillenbrand, J. M., Clark, M. J., & Houde, R. A. (2000). Some effects of duration on vowel recognition. *Journal of the Acoustical Society of America*, 108, 113-122.

Hirata, Y. (2004). Effects of speaking rate on the vowel length distinction in Japanese. *Journal of Phonetics,* 32, 565-589.

Hirata, Y. (2005). Factors affecting native English speakers' identification of Japanese vowel length. *1st ASA Workshop on Second Language Speech Learning*, Vancouver, May 2005.

House, A. S. (1961). On vowel duration in English. *The Journal of the Acoustical Society of America*, 25, 105-113.

Houston, D. M., Santelmann, L. M., & Jusczyk, P. W. (2004). English-learning infants' segmentation of trisyllabic words from fluent speech. *Language and Cognitive Processes*, 19, 97-136.

Ingram, J. C. L., & Park, L.-C. (1991). Cross-language vowel perception and production by Japanese and Korean learners of English. *Journal of Phonetics*, 25, 343-370.

Johnson, E. K., & Jusczyk, P. W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language*, 44, 548-567.

Jones, F. N. (1974). History of psychophysics and judgement. In: E. C. Carterette, & M. P. Friedman (Eds.), *Handbook of Perception.* London: Academic Press.

Jusczyk, P. W. (1986). Toward a model of the development of speech perception. In: J. S. Perkell, & D. H. Klatt (Eds.), *Invariance and Variability in Speech Processes* (pp. 1-35). London: Lawrence Erlbaum Associates.

Jusczyk, P. W. (1997). *The Discovery of Spoken Language.* Cambridge: MIT.

Jusczyk, P. W., & Aslin, D. (1995). Infant's detection of sound patterns of words in fluent speech. *Cognitive Psychology*, 29, 1-23.

Jusczyk, P. W., Cutler, A., & Redanz, N. R. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development*, 64, 675-687.

Jusczyk, P. W., Friedereci, A. D., Wessels, J. M. I., Svenkerud, V. Y., & Jusczyk, A. M. (1993). Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language,* 32, 402-420.

Jusczyk, P. W., & Hohne, E. A. (1997). Infants' memory for spoken words. *Science,* 277, 1984-1986.

Jusczyk, P. W., Hohne, E. A., & Bauman, A. (1999). Infants' sensitivity to allophonic cues for word segmentation. *Perception & Psychophysics*, 61, 1465-1476.

Jusczyk, P. W., Houston, D., & Goodman, M. (1998). Speech perception during the first year. In: A. Slater (Ed.). *Perceptual Development: Visual, Auditory, and Speech Perception in Infancy* (pp. 357-388). Hove: Psychology Press Ltd.

Jusczyk, P. W., Houston, D., & Newsome, M. (1995). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, 39, 159-207.

Jusczyk, P. W., Pisoni, D. B., & Mullennix, J. (1992). Some consequences of stimulus variability on speech processing by 2-month-old infants. *Cognition,* 43, 253-291.

Kemler Nelson, D. G., Jusczyk, P. W., Mandel, D. R. Myers, J., Turk, A., & Gerken, L. (1995). The head-turn preference procedure for testing auditory perception. *Infant Behaviour & Development,* 18, 111-116.

Kondaurova, M. V., & Francis, A. L. (2004). Perception of the English tense/lax vowel contrast by native speakers of Russian. *Journal of the Acoustical Society of America,* 116, 2572.

Kooijman, V., Hagoort, P., & Cutler, A. (2005). Electrophysiological evidence for prelinguistic infants' word recognition in continuous speech. *Cognitive Brain Research*, 24, 109-116.

Kuhl, P. K. (1979). Speech perception in early infancy: Perceptual constancy for spectrally dissimilar vowel categories. *Journal of the Acoustical Society of America*, 66, 1668-1679.

Kuhl, P. K. (1985). Methods in the study of infant speech perception. In: G. Gottlieb, & N. A. Krasnegor (Eds.), *Measurement of Audition and Vision in the First Year of Postnatal Life: A Methodological Overview* (pp. 223-251). Norwood, N.J.: Ablex.

Kuhl, P. K. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics*, 50, 93-107.

Kuhl, P. K., Andruski, J. E., Chistovich, E. R., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., Stolyarova, E. I., Sundberg, U., & Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science*, 277, 684-686.

Kuhl, P. K., Conboy, B. T., Padden, D., Nelson, T., & Pruitt, J. (2005). Early speech perception and later language development: Implications for the "critical period." *Language Learning and Development*, 1, 237-264.

Kuhl, P. K., & Hillenbrand, J. M. Speech perception by young infants: Perceptual constancies for categories based on pitch contour. *Society for Research in Child Development*, San Fransicco, March, 1979.

Kuhl, P. K., & Melzoff, A. N. (1982). The bimodal perception of speech in infancy. *Science*, 218, 1138-1141.

Kuhl, P. K., Tsao, F.-M., & Liu, H.-M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Science USA*, 1996-1201.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science,* 255, 606-608.

Ladefoged, P. (1993). *A Course in Phonetics.* London: Harcourt Brace & Company.

Lieberman, P. (1960). Some acoustic correlates of word stress in American English. *Journal of the Acoustical Society of America*, 4, 451-454.

Liu, H-M., Kuhl, P. K., & Tsao, F-M. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science*, 6:3, F1-F10.

Liu, H., Tsao, F., Stevens, E. B., & Kuhl, P. K. (1999). The relation between infant speech perception and early communicative development: A longitudinal study. *Journal of the Acoustical Society of America*, 106, 2245.

Luce, P. A., & Charles-Luce, J. (1985). Contextual effects on vowel duration, closure duration, and the consonant vowel ratio in speech production. *Journal of the Acoustical Society of America,* 78, 1949-1957.

Mandel, D. R., Jusczyk, P. W., & Kemler Nelson, D. G. (1994). Does segmental prosody help infants organize and remember speech information? *Cognition*, 53, 155-180.

Mattys, S. L., & Jusczyk, P. W. (2001). Phonotactic cues for segmentation of fluent speech by infants. *Cognition*, 78, 91-121.

Mattys, S. L., Jusczyk, P. W., Luce, P. A., & Morgan, J. L. (1999). Word segmentation in infants: How phonotactics and prosody combine. *Cognitive Psychology*, 38, 465-494.

Maurer, D., & Young, R. (1983). Newborns' following of natural and distorted arrangements of facial features. *Infant Behaviour and Development,* 6, 127–131.

Maye, J., Werker, F. J., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, 82, 101-111.

McQueen, J. M. (2005). Speech perception. In K. Lamberts & R. Goldstone (Eds.), *The Handbook of Cognition* (pp. 255-275). London: Sage Publications.

Mehler, J., Jusczyk, P. W., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, 29, 144-178.

Miller, J. L., & Eimas, P. D. (1994). Observations on speech perception, its development, and the search for mechanism. In: J. C. Goodman, & H. C. Nusbaum (Eds.), *The Development of Speech Perception* (pp. 37-55). Cambridge (MA): MIT Press.

Morgan, J. L., & Saffran, J. R. (1995). Emerging integration of sequential and suprasegmental information in preverbal speech segmentation. *Child Development,* 66, 911-936.

Morrison, G. S. (2000). *Effects of L1 Duration Experience on Japanese Listeners' Perception of English Front Vowels.* Unpublished Masters Thesis, Simon Fraser University, Vancouver, Canada.

Morrison, G. S. (2002). *Perception of English /i/ and /ɪ/ by Japanese and Spanish Listeners: Longitudinal Results.* Proceedings of the North West Linguistics Conference 2000. Vancouver: Simon Fraser University Linguistics Graduate Student Association.

Morrison, G. S. (2002). Perception of English /i/ and /I/ by Japanese listeners. In: S. Oh, N. Sawai, K. Shiobara, & R. Wojak (Eds.), *University of British Columbia Working Papers in Linguistics,* Volume 8: *Proceedings of NWLC 2001: Northwest Linguistics Conference* (pp. 113-131). Vancouver, BC: University of British Columbia, Department of Linguistics.

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Towards an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 756-766.

Norris, D. G., McQueen, J. M., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences, 23*, 299-325

Nooteboom, S. G. (1972). Production and Perception of Vowel Duration in Dutch: A study of durational properties of vowels in Dutch. Unpublished doctoral dissertation, University of Utrecht. Also in: Philips Research Reports, 5, (165 pp.).

Nooteboom, S. G., & Doodeman, G. J. N. (1980). Production and perception of vowel length in spoken sentences. *Journal of the Acoustical Society of America*, 67, 276-287.

Oller, D. K. (1973). "The effects of position in utterance on speech segment duration in English." *Journal of the Acoustical Society of America*, 54, 1235-1247.

Pater, J., Stager, C., & Werker, J. F. (1998). Additive effects of phonetic distinctions in word-learning. (manuscript in preparation). In: Fennell, C. T., & Werker, J. F. (2003). Early word learners' ability to access phonetic detail in well-known words. *Language and Speech*, 46, 245-264.

Pater, J., Stager, C., & Werker, J. F. (2004). The lexical acquisition of phonological contrasts. *Language*, 80, 361-379.

Pegg, J. E., & Werker, J. F. (1997). Adult and infant perception of two English phones. *Journal of the Acoustical Society of America*, 102, 3742-3753.

Peterson, G. E., & Lehiste, I. (1961). "Duration of syllable nuclei in English." *Journal of the Acoustical Society of America*, 32, 693-703.

Polka, L., & Werker, J. F., (1994). Developmental changes in non-native perception of vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 431-435.

Ramus, F. (2002). Language discrimination by newborns. Teasing apart phonotactic, rhythmic and intonational cues. *Annual Review of Language Acquisition*, 2, 85-115.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996a). Statistical language learning by 8 month olds. *Science*, 274, 1926-1928.

Saffran, J., Newport, E. L., & Aslin, R. N. (1996b). Word segmentation: The role of distributional cues: *Journal of Memory and Language*, 35, 606-621.

Salverda, A. P., Dahan, D., & McQueen, J. M. (2003). The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension. *Cognition*, 90, 51-89.

Shvachkin, N. K. (1973). The development of phonemic speech perception in early childhood. In: C. Fergusson, & D. Slobin (Eds.), *Studies of Child Language Development* (pp. 91-127). New York: Holt, Rinehart & Winston. (Original work published 1948).

Smits, R., Warner, N., McQueen, J. M., & Cutler, A. (2003). Unfolding of phonetic information over time: a database of Dutch diphone perception. *Journal of the Acoustical Society of America*, 113, 563-74.

Snow, C. E. (1994). Beginning from baby talk: Twenty years of research on input in interaction. In: C. Gallaway & B. J. Richards (Eds.), *Input and Interaction in Language Acquisition* (pp. 3-12). New York: Cambridge University Press.

Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, 388, 381-382.

Strange, W., Akahane-Yamada, R., Kubo, R., Trent, S. A., Nishi, K., & Jenkins, J. J. (1998). Perceptual assimilation of American English vowels by Japanese listeners. *Journal of Phonetics*, 26, 311-344.

Swingley, D. (2001, April). Word recognition in the second year: what develops? Paper presented at the *Biennial Meeting of the Society for Research in Child Development*, Minneapolis.

Swingley, D. (2003). Phonetic detail in the developing lexicon. *Language and Speech*, 46, 265-294.

Swingley, D. (2005a). Statistical clustering and the contents of the infant vocabulary. *Cognitive Psychology*, 50, 86-132.

Swingley, D. (2005b). 11-month-olds' knowledge of how familiar words sound. *Developmental Science*, 8, 432-443.

Swingley, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, 76, 147-166.

Swingley, D., & Aslin, R. N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science,* 13, 480-484.

Swingley, D., Pinto, J. P., & Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition,* 71, 73-108.

Tsao, F.-M., Liu, H.-M., & Kuhl, P. K. (2004). Speech perception in infancy predicts language development in the second year of life: A longitudinal study. *Child Development*, 75, 1067-1084.

Van de Weijer, J. (1999). *Language input for word discovery.* Unpublished doctoral dissertation, University of Nijmegen (MPI Series in Psycholinguistics, 9).

Walker, D. C. (1975). Another idiolect: Comments on an article by Professor Avis. In: J. K. Chambers (Ed.), *Canadian English: Origins and Structures* (pp. 118-128). Toronto: Methuen.

Watson, I. (1983). Cues to the voicing contrast: A survey. *Cambridge Papers in Phonetics and Experimental Linguistics,* 2, 1-34.

Weber, A. (2001). *Language-specific listening: The case of phonetic sequences.* Unpublished doctoral dissertation, University of Nijmegen (MPI Series in Psycholinguisitics, 16).

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language*, 50, 1-25.

Werker, J. F., Cohen, L. B., Lloyed, V. L., Casasola, M., & Stager, C. L. (1998). Acquisition of word-object associations by 14-month old infants. *Developmental Psychology*, 34, 1289-1309.

Werker, J. F., & Fennell, C. T., Corcoran, K. M., & Stager, C. L. (2002). Infants' ability to learn phonetically similar words: Effects of age and vocabulary. *Infancy*, 3, 1-30.

Werker, J. F., & Fennell, C. T. (2004). From listening to sounds to listening to words: Early steps in word learning. In: D. G. Hall, & S. R. Waxman (Eds.), *Weaving a Lexicon* (pp. 79-109). Cambridge, MA: MIT Press.

Werker, J. F., & Lalonde, C. E. (1998). Cross-language speech perception: Initial capabilities and Developmental change. *Developmental Psychology*, 24, 672-683.

Werker, J. F., Pons, F., Dietrich, C., Kajikawa, S., Fais, L., & Amano, S. (submitted). Infant-directed speech supports phonetic category learning in English and Japanese.

Werker, J. F., Shi, R., & Desjardins, R., Pegg, J. E., Polka, L., & Patterson, M. (1998). Three methods for testing infants speech perception. In: A. Slater (Ed.), *Perceptual Development: Visual, Auditory, and Speech Perception in Infancy* (pp. 389-420). Hove: Psychology Press Ltd.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behaviour and Development*, 7, 49-63.

Werker, J. F., & Tees, R. C. (1999). Influences on infant speech processing: Toward a new synthesis. *Annual Review of Psychology*, 50, 509-535.

Werner, L. A. (1996). The development of auditory behavior (or what the anatomists and physiologists have to explain). *Ear and Hearing*, 17, 438-446.

Werner, L. A., & Gray, L. (1998). Behavioral studies of hearing development. In E. W. Rubel, A. N. Popper, & R. R. Fay (Eds.), *Development of the Auditory System*. Vol. 5, *Springer Handbook of Auditory Research* (12-79). New York: Springer-Verlag.

Woodward, Z. N., & Aslin, R. N. (1990). *Segmentation cues in maternal speech to infants.* Paper presented at the International Conference on Infancy Studies, Montreal, Canada.

9 SAMENVATTING EN CONCLUSIES

9.1 Het onderscheiden van contrasterende en niet-constrasterende variatie

In dit proefschrift is onderzocht hoe babies leren welke fonetische variatie wel en niet fonologisch relevant is voor de waarneming van spraak. Hiervoor is een vergelijkend onderzoek verricht naar de verwerking van variaties in klinkerlengte door babies die Nederlands leren en babies die Canadees Engels leren.

Klinkerlengte is fonemisch in het Nederlands maar niet in het Canadees Engels. Nederlandse babies leren daarom dat twee woorden die in uitspraak alleen van elkaar verschillen in de duur van de klinker (bijvoorbeeld "tak" en "taak") twee verschillende woorden zijn, en geen twee varianten van het zelfde woord. Daarentegen leren Canadese babies dat de woorden "cat" en "caat" twee varianten van het zelfde woord zijn. Deze woorden kunnen overigens wel door de babies van elkaar onderscheiden worden omdat het woord "caat" afwijkt van de manier waarop het woord "cat" normaal uitgesproken wordt.

Het is een hele uitdaging voor babies om te leren wat voor rol klinkerlengte speelt in hun moedertaal. De lengte van klinkers wordt beïnvloed door meerdere contextuele factoren, zoals spraaksnelheid, de mate van beklemtoning, de positie van een woord in een zin en de fonetische context van een woord. In het Engels is klinkerlengte bijvoorbeeld een belangrijke aanwijzing voor het bepalen van de stemhebbendheid van de plofklank die op een klinker volgt. De duur van een klinker is langer wanneer deze wordt gevolgd door een stemhebbende plofklank dan wanneer de klinker wordt gevolgd door een stemloze plofklank (bijvoorbeeld "tab" vergeleken met "tap"). Babies die Engels leren kunnen variatie in de duur van klinkers dus niet negeren, omdat klinkerlengte een belangrijke bron van informatie is voor het bepalen van de stemhebbendheid van medeklinkers die volgen op klinkers. Engelse babies moeten echter ook leren dat de lengte van een klinker geen belangrijke informatie vormt voor de herkenning van de klinker en zij moeten dus leren om relatief meer aandacht te besteden aan de kwaliteit van klinkers dan aan verschillen in duur tussen klinkers. In het Nederlands verschillen sommige klinkers hoofdzakelijk in lengte (bijvoorbeeld /ɑ/ en /a/) maar andere klinkers niet (er is bijvoorbeeld wel een korte /ɛ/ maar geen lange /ɛ:/ behalve in bepaalde zeldzame leenwoorden). Nederlandse babies moeten dus leren om variatie in klinkerlengte te gebruiken bij het herkennen van klinkers die verschillen in duur ten opzichte van andere klinkers, terwijl ze ook moeten leren om variatie in klinkerlengte te negeren bij het herkennen van klinkers die niet van andere klinkers verschillen in duur.

In dit proefschrift heb ik onderzocht of babies hebben geleerd dat de lengte van klinkers een rol speelt in de hen omringende taal en als zodanig een belangrijke rol van informatie vormt bij het herkennen van klinkers. Onze resultaten wijzen er op dat babies van 17 maanden inderdaad de fonemische status van klinkerlengte in de hen omringende taal hebben geleerd (zie hiervoor de samenvatting van onze resultaten hieronder).

9.2 De bestendigheid van vroege lexicale representaties voor niet-contrasterende akoestische variatie

In eerder onderzoek is reeds aangetoond dat de klankrepresentaties van babies op het niveau van klanken en op het niveau van lettergrepen taal-specifiek zijn tegen het eind van hun eerste levensjaar. Het merendeel van de studies die de invoed van de moedertaal van babies op de herkenning van woorden hebben onderzocht, hebben daarvoor gekeken naar klankverschillen die fonemische contrasterend zijn. In het onderzoek in dit proefschrift werd daarentegen gebruik gemaakt van een contrast dat fonemisch was voor een groep babies en niet-fonemisch voor de andere groep babies.

Door twee groepen babies te vergelijken die een andere moedertaal hebben, en voor wie klinkerlengte dus ofwel fonemisch ofwel niet fonemisch was, was het mogelijk om verschillende voorspellingen te doen met betrekking tot de verwerking van klinkerlengte tussen de twee groepen babies, gebaseerd op de segmentele fonologie van de moedertaal van de babies. Wanneer Nederlandse en Canadese babies variatie in de duur van klinkers op verschillende manieren verwerken, zou dit een aanwijzing zijn dat de vroege lexicale representaties van babies gedetailleerde informatie over klinkers bevatten. Het zou ook een aanwijzing zijn dat deze representaties taal-specifieke informatie bevatten met betrekking tot aspecten die verband houden met de duur van klinkers. Dit is precies wat we gevonden hebben.

9.3 Omgaan met saillante, niet-fonemische akoestische variatie

Bestaande studies die de verwerking door babies van verschillen tussen klinkers in een andere taal dan de moeder taal hebben bestudeerd, hebben vaak gekeken naar betrekkelijk subtiele verschillen tussen klinkers. In de serie experimenten in dit proefschrift hebben we ons gericht op de vraag hoe babies saillante akoestische variatie in de duur van klinkers verwerken die onderdeel uitmaakt van het spraaksignaal, maar niet fonemisch is voor voor Canadese babies.

Het is bekend dat volwassenen, kinderen en zelfs pasgeborenen gevoelig zijn voor zeer kleine verschillen in duur van stimuli, voor verschillende aspecten van spraakwaarneming (bijvoorbeeld in verband met het compenseren voor spraaksnelheid). Het feit dat zelfs pasgeborenen gevoelig zijn voor deze temporele akoestische variatie wijst er op dat gevoeligheid voor deze bron van variatie wellicht, althans gedeeltelijk, onafhankelijk is van ervaring met taal. Het zou echter ook kunnen dat de akoestische opvallendheid van verschillen in klinkerlengte bijdraagt aan de gevoeligheid van luisteraars voor dergelijke variatie. Het vermogen van luisteraars om subtiele variatie in de duur van stimuli waar te nemen kan voorts verband houden met het belang dat variatie in klinkerlengte in veel talen heeft op verschillende niveau's van verwerking van het spraaksignaal.

In dit proefschrift laten we zien dat de manier waarop luisteraars variatie in de lengte van klinkers interpreteren uiteindelijk wordt bepaald door de hen omringende taal.

9.4 Experimentele resultaten

In Experiment 1 werd de mate van specificiteit van de vroegste lexicale representaties van babies onderzocht met behulp van de Conditioned Headturn Methode (CHT). Nederlandse babies van acht tot twaalf maanden werd geleerd om opzij te kijken wanneer zij een bepaald nieuw woord hoorden. Vervolgens werden deze nieuwe woorden, en controle woorden die op deze woorden leken, in een zinscontext aangeboden en werd gekeken naar de reactie van de babies. De controle woorden verschilden van de nieuwe woorden in de kwaliteit of duur van de klinker. Het verschil in de lengte van de klinker was fonemisch, maar het verschil in de kwaliteit van de klinker was dat niet. De babies keken opzij wanneer zij het nieuwe woord hoorden maar niet wanneer zij een van de controle woorden hoorden, wat er op duidt dat er een toename in activatie van het geleerde (nieuwe) woord was ten opzichte van de (controle) varianten. Deze verschillen in de mate waarin babies hun hoofd draaiden in de experimentele condities waren in dit experiment niet significant, maar de snelheid waarmee de babies hun hoofd draaiden was dat wel. Nederlandse babies reageerden het snelst op de nieuwe woorden. De babies reageerden significant langzamer op het controle woord waarin de klinker van langere duur was, en dus fonemisch verschilde van de klinker in het nieuwe woord, terwijl ze niet significant langzamer reageerden op het controle woord waarin de klinker was vervangen door een klinker die niet in het Nederlands voorkomt. Een manier om dit resultaat te interpreteren is dat dit verband houdt met de segmentele fonologie van het Nederlands. Als de reacties van de Nederlandse babies hun ervaring met de Nederlandse taal weerspiegelen, zou het nieuwe woord, "hasp", voor hen als een ander woord klinken dan het controle woord met de langere klinker, "haasp", terwijl dit niet het geval zou zijn voor het controle woord, "husp". Een alternatieve verklaring voor de resultaten van dit experiment is dat het akoestische verschil tussen de klinker in het nieuwe woord en de klinker in het controle woord groter was voor het controle woord met een klinker die verschilde in kwaliteit ("husp") dan voor het controle woord met een klinker die verschilde in lengte ("haasp").

Om dit te onderzoeken werd in Experiment 2 en 3 een onderscheidingstaak gebruikt waarin Canadese en Nederlandse babies van zeven-en-een-halve tot twaalf maanden oud werden getest. Onderzocht werd of zij, nadat zij hadden geleerd om hun hoofd te draaien bij het horen van een nieuw woord, anders zouden reageren op de verschillende soorten controle woorden in een zinscontext. Hiervoor werd wederom gebruik gemaakt van de CHT methode. De controle woorden [illegible]en van het nieuwe woord in klinkerkwaliteit (dit verschil was fonemisch voor beide groepen babies) of in klinkerlengte (dit verschil was alleen fonemisch voor de Nederlandse babies). Omdat klinkerlengte een verschillende rol speelt in de twee talen, voorspelden we dat Nederlandse babies relatief gezien minder zouden reageren op beide controle woorden, terwijl Canadese babies relatief gezien minder zouden reageren op het controle woord waarin de klinker van kwaliteit verschilde, terwijl zij niet minder zouden reageren op het controle woord waarin de klinker van lengte verschilde. We vonden echter dat beide groepen babies significant minder reageerden op beide soorten controle woorden. Dit wijst er op dat zowel de Nederlandse als de Canadese babies zowel het controle woord met een klinker van verschillende lengte als het controle woord met een kwalitatief verschillende klinker categoriseerden als verschillend van het geleerde (nieuwe) woord.

Het resultaat van het experiment met Canadese babies lijkt verrassend omdat de lengte van een klinker niet fonemisch is in Canadees Engels. Een mogelijke verklaring voor de hierboven beschreven resultaten is dat woorden die in de CHT methode gebruikt worden geen verband houden met objecten. Het is ook mogelijk dat de resultaten zijn beinvloed door de akoestische opvallendheid van klinker lengte. Om de CHT taak goed uit te voeren, moesten de babies een gedetailleerde representatie van het nieuwe woord vormen, deze representatie gedurende de leerfase onthouden en vervolgens akoestisch vergelijken met de test stimuli. Hiervoor was het niet noodzakelijk een verband te leggen tussen een woord en een object. Daardoor konden de babies de taak uitvoeren zonder gebruik te maken van een lexicaal verwerkingsniveau. In plaats hiervan was een lager, akoestisch verwerkinsproces was wellicht voldoende. Dit zou kunnen verklaren waarom Canadese babies gevoelig waren voor variatie in klinkerlengte en dus in staat waren het nieuwe woord te onderscheiden van het controle woord met een klinker die in duur verschilde van de klinker in het nieuwe woord, terwijl de duur van klinkers voor Canadese babies niet fonemisch is. Het is echter ook mogelijk dat de babies wel degelijk een onderscheid maakten op basis van door hen waarneembare perceptuele verschillen, omdat dergelijke verschillen op andere linguïstische niveau's van de interpretatie van spraak noodzakelijk zijn. De duur van een klinker speelt bijvoorbeeld zowel in het Nederlands als het Engels een rol bij het waarnemen van beklemtoning in een woord en het waarnemen van het accent in een zin. In het Engels speelt de duur van een klinker ook een rol bij het categoriseren van de stemhebbendheid van medeklinkers aan het eind van een lettegreep.

Samenvattend hebben we aangetoond dat Nederlandse en Canadese babies in staat zijn om een nieuw woord dat zij geleerd hebben te herkennen in een zin, en dat zij in staat zijn om dit woord te onderscheiden van woorden die verschillen in de kwaliteit of lengte van de klinker. Zelfs Canadese babies, voor wie klinkerlengte niet fonemisch is, bleken dit vermogen te hebben. Omdat informatie met betrekking tot klinkerlengte belangrijk is voor andere niveau's van verwerking in het Engels is, lijkt het aannemelijk dat het belangrijk en zelfs noodzakelijk is voor Canadese babies om gevoelig te zijn voor variatie in de duur van klinkers, omdat dit een eerste vereiste is om doeltreffend te kunnen compenseren voor natuurlijke variatie in spraak (bijvoorbeeld om te compenseren voor spraaksnelheid).

De resultaten van Experiment 2 en 3 riepen de vraag op of taal-specifieke verschillen in de verwerking van de duur van klinkers tussen Nederlandse en Engelse babies zich zou voordoen wanneer oudere babies getest zouden worden in een referentiële taak. Een dergelijke taak zou de Engelstalige babies er toe kunnen bewegen meer gebruik te maken van een lexicaal niveau van verwerking, waardoor zij wellicht variatie in de lengte van klinkers zouden negeren bij het herkennen van klinkers.

In Experiment 4 en 5 werd onderzocht of het gebruik van een meer referentiële taak er toe zou leiden dat taal-specifieke verschillen zouden kunnen worden waargenomen tussen de twee groepen babies, in overeenstemming met de segmentele fonologie van hun taal. We onderzochten of Nederlandse en Canadese babies van 17 maanden oud twee woorden die alleen van elkaar verschilden in de duur van de klinker zouden kunnen leren associeren met twee verschillende objecten, in een zogenaamde "Switch" taak. Wanneer Nederlandse babies wel, maar Canadese babies niet in staat zouden zijn de twee nieuwe woorden te leren, zou dit een sterk bewijs zijn dat de lexicale representaties van babies van 17 maanden gedetailleerde informatie over de duur van klinkers bevatten, in overeenstemming met de segmentele fonologie van hun moedertaal.

De twee groepen babies leerden om twee woorden met een object te associeren. Vervolgens werd gekeken of de babies het zouden merken wanneer de associatie tussen de woorden en de objecten werd omgedraaid.

In tegenstelling tot de vorige studies werd een verschil gevonden tussen de reacties van Nederlandse en Canadese babies met betrekking tot de lengte van de klinkers in de stimuli. Nederlandse babies merkten een verandering in klinkerlengte op, terwijl Canadese babies hiertoe niet in staat waren. In de controle conditie merkten beide groepen babies een verandering in klinkerkwaliteit op (die voor iedere groep relevant was ten opzichte van de moedertaal).

Samenvattend bevestigen deze data de hypothese dat het gebruik van een taak die een beroep doet op een lexicaal niveau van verwerking babies er toe kan aanzetten gebruik te maken van een meer taal-specifieke manier van verwerking, in zoverre dat het Engelstalige babies niet leek op te vallen wanneer de lengte van de klinker in deze meer referentiële "Switch" taak veranderde. Hoewel we niet specifiek hebben getest of Engelstalige babies van 17 maanden het vermogen hebben om woorden die verschillen in klinkerlengte van elkaar te onderscheiden (in Experiment 2 en 3 werden namelijk veel jongere babies getest), is het moeilijk voorstelbaar dat de babies niet gevoelig zouden zijn voor dergelijke variatie, gegeven het feit dat klinkerlengte een belangrijke rol speelt voor andere niveau's van verwerking in zowel het Engels als andere talen, en gegeven de eerdergenoemde publicaties waarin is aangetoond dat pasgeborenen en volwassenen gevoelig zijn voor verschillen in de duur van klinkers.

Experiment 6 en 7 onderzochten twee belangrijke vragen die werden ingegeven door de hierboven besproken experimenten. Allereerst onderzochten we of verschillen in de stimuli tussen de twee talen er toe geleid zouden kunnen hebben dat Canadese babies niet in staat waren om twee nieuwe woorden die alleen van elkaar verschilden in de duur van de klinker ieder te associeren met een bepaald object in de "Switch" taak. De woorden die voor Nederlandse babies waren gebruikt bevatten namelijk veel kortere klinkers dan de woorden die voor Canadese babies waren gebruikt, hoewel de verhouding in de lengte van de klinkers tussen woorden met lange en korte klinkers voor beide groepen 2:1 was. Deze verschillen in de duur van de klinkers vinden hun oorsprong in natuurlijke variatie in de lengte van klinkers in 'infant-directed speech' tussen de twee talen. Het is daarom mogelijk dat de Canadese stimuli, die relatief lang waren, makkelijker van elkaar waren te onderscheiden op grond van akoestische eigenschappen van deze stimuli (bijvoorbeeld omdat een verschil in klinkerlengte gemakkelijker is waar te nemen voor klinkers die relatief lang zijn) en niet als gevolg van verschillen in de segmentele fonologie van de moedertaal van de Canadese babies.

In Experiment 6 werd daarom onderzocht of Nederlandse babies de twee Canadese woorden, die verschilden in klinkerlengte, konden leren associëren met twee nieuwe objecten in de "Switch" taak. Het bleek dat Nederlandse babies hiertoe in staat waren, net zoals zij dat waren met de Nederlandse stimulus woorden. Dit wijst er op dat de natuurlijk verschillen in lengte tussen de twee groepen stimulus woorden in Experiment 3 en 4 niet verantwoordelijk zijn voor het gevonden verschil in de resultaten van deze experimenten.

Experiment 7 werd ingegeven door de vraag of Nederlandse babies klinkerlengte contrasterend zouden verwerken voor alle Nederlandse klinkers, of dat zij dat alleen zouden doen voor klinkers die in het Nederlands kort of lang kunnen zijn. Om dit te onderzoeken werd getest of Nederlandse babies het vermogen hadden om de twee nieuwe woorden "tem" en "teem" te leren. In spraak die gericht is tot Nederlandse babies wordt geen onderscheid gemaakt tussen /ɛ/ en /ɛ:/, hoewel de lange klinker /ɛ:/ wel tot uiting komt in bepaalde zeldzame leenwoorden. Wanneer de babies in staat zouden zijn deze woorden te leren, zou dit een aanwijzing zijn dat het vermogen van babies om verschillen in klinkerlengte op te merken zich niet beperkt tot klinkers in de Nederlandse taal waarvoor klinkerlengte fonemisch is.

De resultaten lieten zien dat Nederlandse babies inderdaad in staat waren de twee woorden te associeren met de twee nieuwe objecten, ondanks het feit dat het verschil in de lengte van de klinker in de Nederlandse taal niet fonemisch is. Dit roept de vraag op of Nederlandse babies een algemene klinker parameter verwerven die hen er toe aanzet om meer aandacht te schenken aan klinkerlengte in het algemeen, en niet alleen wanneer zij klinkers verwerken die in het Nederlands van elkaar verschillen in duur. Het feit dat Canadese babies van 17 maanden niet in staat waren om in de "Switch" taak woorden die van elkaar verschilden in klinkerlengte te leren, wijst er op dat hier inderdaad sprake is van een taal-specifiek effect. Wanneer dit effect namelijk zijn oorsprong zou vinden in akoestische opvallendheid, zouden de Canadese babies in staat moeten zijn geweest om woorden die alleen van elkaar verschilden in klinkerlengte te associëren met nieuwe objecten in de "Switch" taak. Maar dit was niet het geval.

9.5 Conclusie

In dit proefschrift werd het leren van contrasterende en niet-contrasterende akoestische variatie gedurende fonologische verwerving onderzocht. We vergeleken hiervoor de verwerking van variatie in klinkerlengte tussen twee populaties met een verschillende taal, waarin variatie in klinkerlengte wel of niet fonemisch is. Hierbij is belangrijk in gedachten te houden dat de lengte van klinkers in beide talen belangrijke informatie verschaft voor andere niveau's van verwerking dan klinker herkenning. Er werd gebruik gemaakt van een onderscheidingstaak en van een "Switch" taak zodat de verwerking van variatie in klinkerlengte door babies kon worden onderzocht op verschillende niveau's van verwerking en met verschillende taakvereisten.

We hebben aangetoond dat zowel Nederlandse als Canadese babies gevoelig zijn voor verschillen in klinkerlengte wanneer zij tussen de acht en twaalf maanden oud zijn, en dat zij van deze gevoeligheid gebruik kunnen maken in onderscheidingstaken.

Op de leeftijd van 17 maanden was er een verschil ontstaan tussen de twee groepen babies, in die zin dat wanneer zij getest werden in een "Switch" taak, Nederlandse babies gebruik konden maken van variatie in klinkerlengte tijdens het leren van de woorden maar Canadese babies niet.

Samenvattend zijn deze resultaten in overeenstemming met de segmentele fonologie van het Nederlands en van het Canadees-Engels. Het onderscheiden van stimuli op grond van klinkerlengte lijkt belangrijk te zijn voor zowel Nederlandse als Engelse babies, in beide leeftijdsgroepen, omdat informatie met betrekking tot klinkerlengte in beide talen van belang is voor andere aspecten van verwerking. Daarentegen is aandacht schenken aan de duur van klinkers bij het leren van woorden wel belangrijk voor Nederlandse babies, maar niet voor Canadese babies. Onze data weerspiegelen deze beide bekwaamheden. Het lijkt er dus op dat de kennis van fonologische structuur van babies van 17 maanden reeds gespecificeerd is met betrekking tot klinkerlengte, wat in overeenstemming is met de segmentele fonologie van de respectievelijke talen.

CURRICULUM VITAE

Christiane Dietrich was born in Aachen, Germany on April 25^{th}, 1974. After graduating from Einhardt-Gymnasium in Aachen in 1993, she moved to the United Kingdom where she worked as an au-pair and a nanny. Christiane subsequently completed a four-year B.Sc. honours degree in Speech Sciences at University College London and obtained a license to practise as a speech and language therapist.

In November, 1999 Christiane was awarded a scholarship by the German Max-Planck-Institute for Psycholinguisitics (MPI). During her doctoral studies at the MPI, Christiane had the opportunity to visit Dr. Janet Werker's Infant Studies Centre at the University of British Columbia (UBC) for nine months, where she collected data for her research as well as working as a laboratory coordinator of Janet Werker's newborn speech perception laboratory at the BC Children's & Women's Hospital.

After completing her doctoral research, Christiane was awarded a postdoctoral scholarship by NTT Laboratories in Kyoto, Japan to work on a collaborative research project between NTT-Labs and Dr. Werker's Infant Studies Centre at UBC, which is where Christiane was based.

From September, 2004 until August, 2005, Christiane worked towards an M.Sc. in Speech Sciences at UBC and became officially certified to work as a speech-language pathologist (SL-P) in Canada in December, 2005. Christiane has been working as an SL-P at G.F. Strong Rehabilitation Centre in Vancouver, British Columbia since that time, providing swallowing and communication assessment and rehabilitation services to adult clients within the 'Acquired Brain Injury Inpatient Program'.

MPI SERIES IN PSYCHOLINGUISTICS

---

1. The electrophysiology of speaking: Investigations on the time course of semantic, syntactic, and phonological processing. *Miranda van Turennout*
2. The role of the syllable in speech production: Evidence from lexical statistics, metalinguistics, masked priming, and electromagnetic midsaggital articulography. *Niels O. Schiller*
3. Lexical access in the production of ellipsis and pronouns. *Bernadette M. Schmitt*
4. The open-/closed-class distinction in spoken-word recognition. *Alette Haveman*
5. The acquisition of phonetic categories in young infants: A self-organising artificial neural network approach. *Kay Behnke*
6. Gesture and speech production. *Jan-Peter de Ruiter*
7. Comparative intonational phonology: English and German. *Esther Grabe*
8. Finiteness in adult and child German. *Ingeborg Lasser*
9. Language input for word discovery. *Joost van de Weijer*
10. Inherent complement verbs revisited: Towards an understanding of argument structure in Ewe. *James Essegbey*
11. Producing past and plural inflections. *Dirk Janssen*
12. Valence and transitivity in Saliba: An Oceanic language of Papua New Guinea. *Anna Margetts*
13. From speech to words. *Arie van der Lugt*
14. Simple and complex verbs in Jaminjung: A study of event categorisation in an Australian language. *Eva Schultze-Berndt*
15. Interpreting indefinites: An experimental study of children's language comprehension. *Irene Krämer*
16. Language-specific listening: The case of phonetic sequences. Andrea Weber

...ng eyes and naming objects. *Femke van der Meulen*

18. Analogy in morphology: The selection of linking elements in Dutch compounds. *Andrea Krott*
19. Morphology in speech comprehension. *Kerstin Mauth*
20. Morphological families in the mental lexicon. *Nivja H. de Jong*
21. Fixed expressions and the production of idioms. *Simone A. Sprenger*
22. The grammatical coding of postural semantics in Goemai (a West Chadic language of Nigeria). *Birgit Hellwig*
23. Paradigmatic structures in morphological processing: Computational and cross-linguistic experimental studies. *Fermín Moscoso del Prado Martín*
24. Contextual influences on spoken-word processing: An electrophysiological approach. *Daniëlle van den Brink*
25. Perceptual relevance of prevoicing in Dutch. *Petra M. van Alphen*
26. Syllables in speech production: Effects of syllable preparation and syllable frequency. *Joana Cholin*
27. Producing complex spoken numerals for time and space. *Marjolein Meeuwissen*
28. Morphology in auditory lexical processing: Sensitivity to fine phonetic detail and insensitivity to suffix reduction. *Rachèl J. J. K. Kemps*
29. At the same time...: The expression of simultaneity in learner varieties. *Barbara Schmiedtová*
30. A grammar of Jalonke argument structure. *Friederike Lüpke*
31. Agrammatic comprehension: An electrophysiological approach. *Marlies Wassenaar*
32. The structure and use of shape-based noun classes in Miraña (North West Amazon). *Frank Seifart*
33. Prosodically-conditioned detail in the recognition of spoken words. *Anne Pier Salverda*