THE PERCEPTION OF NON-NATIVE DIALECTS AMONG CHILDREN AND ADULTS

Joanne Kang
Linguistics Senior Thesis
Fall 2004

INTRODUCTION

Interactions involving different dialects are commonplace due to the constant flow of people in and out of various communities. Children and adults are required to adapt to the new dialects in order to understand what is going on around them. From birth, children have the incredible potential to learn any phoneme that exists in human languages. However, research has shown that as children are exposed to more of their native language, they gradually learn which phonemes are available for use in their language's phonological system and which are not. Even before adulthood, there is an established categorization of one's language sounds. According to Best's (1994) perceptual assimilation model, when people are exposed to non-native dialects, they will categorize the new non-native speech sounds to the closest existing phonological category based on their mental representation of the respective native language.

Using the perceptual assimilation model, I will do a comparative study of how adults and children process non-native dialects. To measure the degree of understanding, I will look at results from studies done on adults and on children where the subjects are asked to reproduce and translate speech of unfamiliar dialects.

BACKGROUND

Dialect

Chambers and Trudgill define *accent* as "the way in which a speaker pronounces and…refers to a variety which is phonetically and/or phonologically different from other varieties" (1980: 5). *Dialect* is defined as "varieties which are grammatically (and perhaps lexically) as well as phonologically different from other varieties" (1980: 5).

However, Chambers and Trudgill point out that linguists use these terms interchangeably because they are very much related to one another and can easily "merge into one another without any discrete break" (1980: 5).

J. C. Wells (1982, cited in Nathan, Wells & Donlan 1998) presents four ways that dialects differ.

The first is that dialects differ in how they are realized phonetically, such as how one vowel can be represented by different phonemes in separate dialects.

Example: The vowel in *game* can be realized as [e] or [ɛɪ] in the Glaswegian and London dialects, respectively.

The second is that dialects can differ on a "structural level in their phonotactic distribution," the environment in which a certain phoneme can exist (Nathan, Donlan & Wells 1998: 344).

Example: /r/ is rhotic in the Glaswegian dialect, able to appear in a variety of environments, while in the London dialect it cannot appear before a consonant or in absolute-final position.

The third way dialects can differ is in the phoneme, or the number of phonemes that can be used to represent a certain vowel combination.

Example: For the words *boot* and *foot*, Glaswegian has only one representation of /ʏ/ while in most other English accents, the words are represented by two separate phonemes of /u/ and /ʊ/.

Finally, dialects differ in the lexical distribution, or how dialects use different phonemes depending on the lexical item.

Example: The word bath is represented as /ɑ/ and /æ/ in southern and northern England dialects, respectively. This is in spite of the fact that both dialect make use of each of the two phonemes elsewhere, such as in *cat* and *cart*.

This paper is concerned only with this type of phonetic variance within the English language. Throughout the rest of this paper, these phonetic differences will be referred to as differences in *dialect*.

Sound Discrimination and Age

From infancy, children have the ability to learn any phoneme that exists in the human language. Research has shown that children can distinguish between native and non-native speech contrasts at as early an age as 6 to 8 months. Within the next few months of infancy, children begin to lose this ability and by adulthood discrimination of non-native speech contrasts becomes increasingly difficult.

Child researchers use several techniques such as the High Amplitude Sucking (HAS) Procedure and the Conditioned Head-Turn (HT) Procedure to study infant speech perception. HAS is useful for experimentation on infants 6 months old or younger. It involves pairing a loud sound stimulus with an image projected on a blank wall in front of an infant. The infant sucks on a nipple that is connected by rubber hose to a pressure transducer. A polygraph machine is used to measure the sucking. The change in sucking rate is correlated with the infants' sensitivity to sound changes (Polka, Jusczyk and Rvachew 1995).

The Conditioned Head-Turn technique (HT) is used more with infants ages 5 months to 18 months old. For this process, the infant sits on the lap of a parent and faces an assistant. The assistant tries to keep the infant's attention by showing various silent toys. On the other side of the infant is a machine that creates a sound stimulus and a dark plexiglass box with a toy animal concealed inside. When the box is activated, it lights up allowing the toy to be visible A constant sound stimulus plays at intervals with occasional changes to the sound for brief periods. Initially the sound and visual stimulus are triggered together, but gradually they grow farther apart in time. Then without a visual stimulus, a change in the sound stimulus is presented. Whether or not the infant responds by turning his or her head in reaction to the sound is the criteria for sound discrimination (Polka, Jusczyk and Rvachew 1995).

Werker and Tees (1983, cited in Werker 1995) first conducted a HT experiment with 6 to 8 month old English-learning infants. They studied the infants' ability to discriminate between Hindi consonant contrasts. The results from that were compared with the results of testing the ability of English-speaking adults and Hindi-speaking adults in hearing the same Hindi consonant contrasts. Results showed that while the infants and adult Hindi speakers could hear the contrasts, the English-speaking adults could not.

A later study by Werker and Tees (1984, cited in Werker 1995) compared English-learning 6 to 8 month old infants and 10 to 12 month old infants in their ability to discriminate both a Hindi consonant contrast and an Interior Salish (Nthlakampx) consonant contrast. The 6 to 8 month old infants heard the sound contrasts while the 10 to 12 month old infants could not.

Children fine-tune their phonological system through exposure to variation. As children hear more sounds that are native to their language, a systematic categorization of the store of native phonemes arises. The absence of non-native sounds does not allow for those sounds to enter into the phonological categorization.

Werker and Pegg (1992) found that it isn't until a child is around 19 months old that a discrimination of words based on minimal pair contrasts was possible, even though the non-native sound discrimination is available 7 to 9 months earlier. To explain this, Werker and Pegg hypothesized that the ability to discriminate the words came from what they called "language-specific phonetic" perception. This term is not to mean that one is limited to hearing the speech contrasts of one's native language alone, but that there is a change in sensitivity to the phonetics of the native dialect.

Before the 10 to 12 month period, infants are using "language-general phonetic" perception. In this stage, infants are able to hear both native and non-native contrasts. At 10 to 12 months, "language-specific phonetic" boundaries are set, but the infants are still unable to distinguish words, which doesn't happen until later. Therefore, "language-specific" perception is the first step in acquiring a phonological categorization and eventually leads to the ability to discriminate between words.

Speech Perception & Environment

There are several theories on how the development of speech perception and environment are related. Pisoni, Lively, & Logan (1994) outline four of these theories.

General Universal Theory:

The general universal theory assumes that infants have the ability to fully perceive any possible phonemic contrasts in language. The perceptual ability is present at birth but the child needs specific types of early experiences to **maintain** the ability. Absence of such experiences could result in partial or complete loss of perceptual ability to discriminate certain speech sounds.

Attunement Theory:

The attunement theory assumes that perceptual ability for speech sounds is only partially developed at birth, inferring that infants can only distinguish some sounds. Specific types of early experiences are required to **facilitate** further development into the full capacity of perception. In the absence of such experiences, there is an absence of further development or a loss in perception, compared to birth.

Perceptual Learning Theory:

The perceptual learning theory assumes that perceptual ability is absent at birth and that development depends on a process of **induction** based on early experiences in their speech environment. Therefore, those specific experiences are necessary for subsequent development and maintenance of perceptual ability.

Maturational Theory:

The maturational theory assumes that early experiences have no role in the development of perceptual ability. Rather, a child's perceptual ability develops on a predetermined developmental schedule. So the age at which specific phonetic contrasts can be made depends on the relative development of the child's sensory

mechanisms. It makes no claims to whether or not perceptual ability is present or absent at birth.

Pisoni, Lively, & Logan suggest that it is most likely a "hybrid of the theories" that accounts for speech development (1994: 127). Also, while the first three theories seem to belong to the same category, the fourth is orthogonal to the rest. It is quite possible that once a person reaches the developmental stage for speech perception, one of the first three theories presented could come into play. Wherever the initial point in development is, it could be assumed that the environment could have an effect from that point on. It does not make sense to assume that experience and exposure to languages do not have an effect on speech perception.

Speech Perception Theories

There are three main theories of speech perception: psychoacoustic theory, motor theory, and direct realist (or ecological) theory. The psychoacoustic theory says that the source of information for speech perception is based on the "*proximal stimulus*, or the raw acoustic components into which the speech signal is assumed to be decomposed by the auditory periphery" (Best 1994: 175). This refers to auditory cues such as waveform characteristics. The psychoacoustic theory assumes that the "perceptual primitives" of speech are meaningless acoustic features, such as "spectra distribution patterns, bursts of bandlimited aperiodic noise, and temporally defined silent gaps" (Best 1994: 175).

Both the motor theory and direct realist theory assume that the perceptual primitives of speech are physical articulatory movements or "gestures," instead of

acoustic cues. These two theories assume that there does not need to be an analysis of the acoustic features because the gestural information can be directly interpreted from the speech signal. Rather than recognizing waveforms and such, there is a recognition of gestures involved in speech like glottal opening, bilabial closure or velum opening.

The direct realist theory assumes that these articulatory gestures are produced by a speaker in the vocal tract. From there, the information is "*directly* detected in speech…*not* built up from an analysis of simple acoustics features" (Best 1995: 177). In other words, if a person hears a phoneme that is unfamiliar to them, the perception of the sound will depend on what the person believes are the necessary articulatory movements needed to produce that sound. When the phoneme is correctly matched with the gestures, the person can then perceive the sound. This does not mean that a person must first produce the sound in order to perceive it. It is only referring to the knowledge of which gestures are needed to be carried out in production of the sound.

The motor theory differs slightly by stating that the gestures are not physical, but rather a representation of "intent" of gesture that exists in the mind of the speaker (Best 1995). Unlike the direct realist theory, the motor theory assumes that an innate "module" in the mind mediates the articulatory gestures. Instead of a direct access to the articulatory gestures like the direct realist theory, the motor theory assumes that the gestural information must first be analyzed in the mind for the neuromotor commands that would bring about such gestures.

Another point of disagreement between the three theories is how the perceptual primitives become developmentally related to morphemes and other linguistic elements. The psychoacoustic theory assumes that infants learn to associate combinations of the

meaningless acoustic features with these linguistic elements. Therefore, infants form "auditory mental traces…that are paired-associates of abstract linguistic entities" through experience (Best 1995: 177).

On the other hand, both the direct realist theory and motor theory state that infants must figure out which articulatory gestures have linguistic meaning in their native languages. The direct realist theory says that gestures are initially void of linguistic meaning and that infants must eventually discover relationships between "higher-order invariants of relations among gestures" and linguistic elements (Best 1995:178).

The motor theory does not assume that the gestures are initially void of linguistic meaning, but that every piece of information the infant receives in speech *is* linguistic by nature. This happens because there is a different mechanism that processes nonlinguistic sound perception.

The direct realist approach says that infants perceive gestures "via an integrated general perceptual system that detects information about distal articulatory events" (Best 1995: 178). Infants learn that speech is used as a communicatory tool within language, and this understanding is central to understanding speech perception. In other words, infants perceive speech in terms of its linguistic goals, meaning perception and production of speech are "inextricably linked" (Best 1995: 179).

The motor theory assumes that a "biologically specialized module relates the incoming speech signal to abstract phonological units via the neuromotor representations of intended phonetic gestures" and it translates the information into "neuromotor commands for producing specific utterances (Best 1995: 178). Therefore, perception and production of language are linked directly. However, studies have shown that people can

perceive non-native dialects without having the ability to produce it. So contrary to the motor theory, there can be perception without production, meaning the two processes are not directly linked.

The following chart of the three theories is taken directly from Best (1995: 176):

Models			
	Psychoacoustic	Direct realist	Motor theory
Assumptions			
Perceptual primitives	Proximal acoustic cues	Distal articulatory gestures	Speaker's *intended* gestures (neuromotor commands)
Perceptual philosophy	Indirect: information processing or mental representation	Direct pick-up of distal; information	Indirect, via motor representations
Perceptual mechanisms	Basic auditory system, aided by cognitive processes	Integrated perceptual systems and their exploratory activities	Specialized phonetic processes of the language module
Specificity re: human speech	General across nonspeech and across other species	General across nonspeech and across other species	Specific to speech and to humans
Relation between perception and production	Not addressed (presumably mediated by cognitive processes)	All linguistic systems integrate perceiving and acting: affordances	Single, specialized module is the source of parity between perception and production
Information infants initially perceive	Nonlinguistic (auditory)	Nonlinguistic (gestural)	Linguistic
Effect of language experience	Formation of traces, templates, prototypes	Perceptual attunement economizes pick-up of native gestural invariants	Native phonetic input tunes the speech module

While these are the three widely known theories about speech perception, it is possible that there could be many more. Some of the data presented above appear to be

orthogonal and could be reorganized under the heading of a different theory, resulting in a "mix-and match" theory.

Best's Perceptual Assimilation Model

Best presents a model for how listeners process non-native speech sounds, called the Perceptual Assimilation Model. Best takes a direct realist approach, stating that listeners "perceive in nonnative phones information about their gestural similarities to native phonemes" (Best 1994: 190). The main idea behind the model is that non-native phonemes are usually recognized for their similarities to the closest phonological representations in the listener's native dialect. In the opposite case where the listener perceives a big discrepancy and cannot assimilate the sound into any phonological category, it may be classified as a nonspeech sound. The following points taken directly from Best (1995: 194) outline the assimilation of non-native speech sounds according to this model:

1. *Assimilated to a native category:* clearly assimilated to a particular native segmental category, or perhaps to a cluster or string, in which case it may be heard as:
    a. a good exemplar of that category
    b. an acceptable but not ideal exemplar of the category
    c. a notably deviant exemplar of the category
2. *Assimilated as uncategorizable speech sound:* assimilated within native phonological space as a speechlike gestural constellation, but not as a clear

exemplar of any particular native category (i.e., it falls within native phonological space but in between specific categories)

3. *Not assimilated to speech (nonspeech sound):* not assimilated into native phonological space at all; heard, instead, as some sort of nonspeech sound

Non-native speech contrasts are assimilated from a similar method as the individual phonemes. Best summarizes the assimilation of contrasts as follows:

Two-Category Assimilation (TC Type)

The phonemes of a non-native contrast may be similar to two corresponding native phonemes. Each non-native phoneme is then assimilated separately to a different category. For example, the Hindi retroflex stop /ɖ̪/ will probably assimilate to the English [d] while the contrasting Hindi breathy-voiced dental stop /ɖ̪^{h}/ will probably assimilate to the English voiced-dental fricative [ð]. Discrimination of these non-native contrasts is expected to be excellent.

Category-Goodness Difference (CG Type)

The phonemes of a non-native contrast are assimilated to one native category, but they differ in how well they fit the native "ideal." One phoneme can be more similar than the other as it is with the Zulu voiceless-aspirated velar /k/ and ejective /k'/. They are both assimilated to English [k^{h}], but the first phoneme is perceived as being a very good match while the second is a little deviant. Discrimination between the non-native contrasts is expected to be moderate to very good, depending on how different the phonemes in the non-native contrast are from each other.

Single-Category Assimilation (SC Type)

The phonemes of a non-native contrast are both assimilated to the same native category, but are both equally different or acceptable from the native "ideal". Such a case is the Thompson Salish ejective velar /k'/ and uvular /q'/. This contrast is assimilated to the English [k^h] even though both do not fit the phoneme perfectly. Discrimination between contrasts is expected to be poor because the two phonemes are equally bad or equally good.

Both Uncategorizable (UU Type)

The sounds of a non-native contrast fall closely "within phonetic space, but outside of any particular native category" (Best 1995:195). Discrimination can range from being poor to very good. Discrimination depends on how similar the sounds in the contrast are to each other and how similar they are "to native categories within native phonological space" (Best 1995:195).

Uncategorized versus Categorized (UC Type)

The sounds of a non-native contrast are split in this case. One non-native sound is assimilated to a native category, and the other is not because it falls outside native categories. Discrimination between the sounds is expected to be very good.

Nonassimilable (NA Type)

The sounds of these non-native contrasts do not fit any speech sound. They are heard rather as nonspeech sounds. For example, the suction-produced click consonants of the southern Bantu languages do not assimilate to any English phoneme. Discrimination is expected to be good to very good depending on their distinctiveness as non-speech sounds.

CHILD DATA

Child Study 1

J. K. Chambers (1988) conducted a study where he studied dialect acquisition among six children of two Canadian families who had moved to Oxfordshire in southern England. The children were of age 9, 13, 14, 15 and 17, two of the children being 13 years old. The children were interviewed twice, with two years separating the interviews. The interviews consisted of a discussion of the children's attitudes toward their new and old neighborhoods, evaluation of tapes of people speaking, identification of objects on cards, and reading of word lists. Chambers specifically looked for changes in the application of certain Canadian dialect features: t-voicing, merging of two low back vowels to one, absence of the southern English vowel backing, presence of postvocalic /r/, and absence of the southern English [r] insertion.

In almost 100% of the 9 year old's words, t-voicing was dropped like it is in the southern English dialect. The child had completely assimilated into the English dialect on this regard. For one 13 year old and a 14 year old, they showed an absence of t-voicing 80% and 90% of the time, respectively. For the other 13 year old, 15 year old, and 17 year old, they only deleted the t-voicing in 20% of their words.

For the absence of the Canadian low back vowel merger, the 9 year old and a 13 year old had the highest percentages of 90% and 80 % respectively. The rest of the children, showed an absence of the low back vowel merger in only 10% or less of the words.

A presence of the vowel backing feature of southern England was evident in 100% of the 9 year old's words, while the rest of the children showed a 20% or less frequency of the vowel backing.

The deletion of postvocalic /r/ was only evident in 10% or less by the majority of the children. However, the 14 year old deleted the /r/ in 30% of his words.

The presence of [r] where there is no lexical /r/ was shown in 40% of the 9 year old's words while the rest of the children did not incorporate this feature of the southern English dialect into any of their words.

Chambers (1992) presents his results as support for eight principles of dialect acquisition.

Lexical replacements happen faster than phonological ones

The identification of objects in pictures showed that the children readily replaced their Canadian lexical items with English lexical items. Examples of such lexical replacements are *coach* for *bus*, *dustbin* for *garbage can*, and *jumper* for *sweater*. But even with these lexical replacements, the children were far more likely to continue to pronounce the new lexical words in their old Canadian dialect. The general trend among the ages was that the youngest children replaced words the most and the oldest replaced words the least.

Lexical replacements happen faster in the first stages of dialect acquisition than in the latter stages

When the children were tested again two years later, it was found that there was only a small increase in the lexical items taken on over the two years. This shows that

most lexical replacements happen within the first two years and anything not replaced is likely to remain unchanged.

Simple phonological rules change faster than more complex ones

Chambers defines "simple phonological rules" as ones that have no exceptions while "complex phonological rules" can have exceptions in form. This refers to exceptions in lexical rules, which depend on the environment of the phonological feature. The children were more likely to drop their t-voicing to a greater degree than with taking on vowel backing. T-voicing is a simple rule in which a /t/ is voiced to a [d] when it follows a vowel or /r/ and comes before an unstressed syllable, such as in the words *putting* vs. *pudding* and *hearty* vs. *hardy*. Vowel backing is a complex rule of lengthening and backing of short /a/ before voiceless anterior fricatives and before clusters of /n/ + obstruent, such as in the words *dancing*, *branch*, and *transmission* whose vowels are pronounced as [ɑ]. Exceptions to this rule, even though the environments are as prescribed, are words like *cafeteria*, *classic*, *pants* and *cancer* whose vowels are pronounced as [æ].

Acquisition of complex rules and new phonemic contrasts distinguishes early acquirers from later acquirers.

The general trend between ages was that the youngest children acquired more dialectal characteristics of their new environments than the older children. For example, while the children acquired the complex rule of vowel backing, in order to lose the complex rule of low vowel merging in Canadian English, the children had to learn new phonemic contrasts. In Canadian English, *tot* and *taught* are pronounced the same, but they are pronounced with two different phonemes in the southern

English dialect. The results showed the youngest children were better at learning the /ɒ/ and /ɔː/ contrast than the older children who ultimately kept the vowel merging.

In early acquisition, "both categorical rules and variable rules of the new dialect result in variability in the acquirers" - Categorical rules are phonological rules which all speakers of a certain dialect display while variable rules are rules that do not have as widespread usage among a dialect group

Simply put, new dialect features cannot be acquired without interfering with the old dialect features. This cannot happen in one sudden change, rather the loss of dialect features and presence of new dialect features is acquired gradually. Dialect features such as t-voicing, which is a categorical rule for Canadian English, were lost variably, occurring only in certain words, and features like vowel backing were acquired variably by the children.

Phonological changes are brought about from pronunciation variants

Chambers says that only after a "critical mass" of pronunciation changes is acquired that it becomes a phonological change. For example, in the case of the acquisition of the /r/ deletion of the southern English dialect, it is only after a child has encountered several words that have the r-less construction that he or she can incorporate the /r/ deletion as a phonological rule and learn to apply the rule. It can be said that acquisition of a dialect happens on a "word by word" basis.

Loss of old rules happens faster and more readily than acquisition of new rules

In the data for the children, there is a greater percentage of the absence of Canadian dialect features than there is of the presence of Southern England dialect features. Chambers (1988) provides a table showing the data.

Phonological Feature	Group Score (%)
Absence of t-voicing	55
Absence of vowel merger	31.6
Presence of vowel backing	23.3
Presence of r-lessness	8.3
Presence of intrusive [r]	6.6

As the table shows, the loss of the Canadian dialect features of t-voicing and low back vowel merging occurs with a much greater percentage than the acquisition of the new southern English dialect rules like /r/ deletion and intrusive [r].

"Orthographically distinct variants are acquired faster than orthographically obscure ones"

In cases like t-voicing, one can see how easy it is to change from a reading of /t/ with a voiced [d] to a reading of /t/ as voiceless [t] since it is present in the words themselves, hence "orthographically distinct." R-dropping is "orthographically obscure" because although the /r/ is present in writing, the dialect requires that it is absent in pronunciation.

Chambers' study shows how age plays a factor in the acquisition of non-native dialects. The fact the younger children, or more specifically the youngest child, acquired so many of the southern English dialect features while the older children did not is probably because the older children have passed the critical age for learning new dialects. It is safe to assume that as adults, it would be even harder to acquire and adapt a new non-native dialect.

Perhaps there is another explanation as well. Although no one will doubt that "ability" to acquire new dialects plays a large role, there may be a motivational aspect to

this as well. It may be that the younger children are more willing to accept new non-native dialects. Based on observation, younger children seem more open to change, especially when it comes to a change in lifestyle and environment. Older children, being more set in their ways, may not have the motivation to learn and adapt to their new environments and might just be content with how they are.

Child Study 2

Nathan, Wells, and Donlan (1998) conducted a study where they hypothesized that children accommodate new and different dialects by making changes to their developing store of phonological representations, as in Best's (1994) Perceptual Assimilation Model. The subjects for this study were children ages four and seven who spoke in the London dialect. The non-native dialect used was the dialect of Glasgow. None of the participating children had any previous contact or experience with the Glaswegian dialect.

The study involved playing a recording of two separate word lists of 20 words each. One word list was read in a Glaswegian dialect and the other word list was read in the familiar London dialect. There was a second recording made with the same words read in the alternate dialects, so the words read in the Glaswegian dialect on the first recording were read in the London dialect on the other tape recording and vice versa. The children were divided into two groups of four year olds and two groups of seven year olds. One group of four year olds and one group of seven year olds listened to one version of the word lists, and the other two groups heard the alternate version of the word lists.

The subjects were asked what they thought the word was as to elicit a repetition of the word, not just an imitation of what they had heard. They were also asked to provide the definitions for the words on the lists.

The children's responses fell into four different categories: phonological repetition, phonetic response, lexical error response, and no response.

- A phonological response was when the child repeated the given word correctly in his/her own dialect. This means that if the child was given the word *church* in the Glaswegian dialect, they would respond back "church" in their London dialect and assign it the definition of "some people go to marry" or something similar.
- A phonetic response was a response where a child repeated the given word phonetically whether he/she knew the true meaning of the word or not. Using the *church* example, if given that word, the child would repeat the word exactly how he/she heard it. The child could possibly define the word (1) correctly as "a place where people go to marry"; (2) incorrectly as a similar sounding word in the London dialect such as the meaning for the word *touch*, "when you put you hand on something"; (3) incorrectly as a word completely unrelated nor similar sounding to the original as "eating"; or (4) give no response at all.
- A lexical error response would be one where a child would produce a similar sounding word in their own London dialect. So given *church* the response would be "touch" said in their London dialect and defined as "when you put your hand on something."
- A "no response" was when a child did not respond with a repetition and definition for the given word.

The study found that older children were far better when it came to repeating back the correct Glaswegian word, pronounced in their London dialect, and assigning it the correct corresponding meaning. These subjects provided the most correct phonological responses.

A type of mistake that the children made was to phonetically repeat back what they heard in the Glaswegian dialect. The younger children proved better at imitating the words than the older children, either that or the imitation was a "default" for the younger children if they could not understand what the corresponding word in their native dialect was.

Both groups performed similarly in producing the wrong lexical response, interpreting the Glaswegian form of the word for another in the London dialect (given "church," repeated back the word "touch" and assigning meaning of "feeling"). The following table shows the distribution of results.

Type of Error	Age 4 (%)	Age 7 (%)
Phonological response	37.0	70
Lexical error repetition	18.0	24
Phonetic response	44.0	4.8
No response	1.5	0.6

The fact that the older children did better at correctly giving a phonological response shows that with age, making sense of non-native dialects improves probably due to a greater exposure to dialect differences. In terms of production, the younger children were better at imitating the non-native dialect than the older children. Nathan, Donlan and Wells believe that this shows that younger children are better at accommodating the

new dialect into their own speech and have a better potential to learn and adapt the dialect as their own.

Regarding the lexical mistakes, Nathan, Wells and Donlan (1998) refer to Labov's (1989) experiment with adults in Chicago and Philadelphia which will be discussed in further detail later on. They hypothesized that the older children would have made less lexical mistakes if the words on the word list were given to them with contextual information.

It would be interesting if the data were arranged a bit differently. For "phonetic responses," whether the child provided a correct or incorrect definition did not come into consideration. The fact that the child repeated the word back in the non-native dialect was the only criteria. If the data for children who repeated the word phonetically with the correct definition and the data for children who repeated the word with the incorrect definition were separate, one could possibly take this study a little further.

It is safe to assume that if an older child recognized a word presented to him or her, the word was repeated back in a London dialect. This being from the fact that 70% of the older children produced a phonological response and only 4.8% produced a phonetic response. Either way, whether the phonetic response was accompanied by a correct or incorrect definition, the data shows that the older children preferred to reproduce the word in their own dialect.

With the younger children, this assumption of a preference for simple word repetition is unclear. Although only 37% produced a phonological response, 44% produced a phonetic response. It would be interesting to see what percentage of the 44% correctly identified the meaning of a word, but decided to repeat it back in the

Glaswegian dialect instead of the London dialect. If there were a great percentage of children who did this, one could assume that there was some relation between age and preference or desire to produce new dialects. For example, while the majority of older children who knew the correct definition of the word decided to repeat the word in their own native dialect, the younger children could have known the correct definition but decided that instead of repeating the word in their native dialect, imitated the dialect that they heard.

Perhaps at the younger age, the children are still be trying to figure out a "dialect identity" for themselves. It could be that the children have not consciously decided what exactly their native dialect is and they are still in the process of figuring that out by experimenting with other dialects and such. Or it could be, perhaps like in the Chambers (1988) study, that the younger children are just more willing to accommodate new dialects into their own dialect.

ANALYSIS OF CHILD DATA

In both child studies, it is clear to see that age is an important factor in determining how well the children produce the sounds of non-native dialects. For both the Chambers (1988) and Nathan, Wells and Donlan (1998) studies, the trend is that the younger the child is when he/she is exposed to a new dialect, the better he/she is at reproducing it. But as discussed earlier, production and perception are not directly related.

In the Nathan, Wells and Donlan (1998) study, this supposed proficiency in the younger group for the production of new phonemes is accompanied by evidence of a lack of understanding of the data presented. The younger age group was not as successful as

the older group in correctly matching the semantic meaning of a word with the non-native word stimulus. Within this age range, the understanding of non-native dialects increases with age.

These results follow Best's perceptual assimilation model, that unfamiliar sounds are assimilated to a similar category in native speech. The fact that the children gave wrong lexical matches to words shows that they were in fact hearing the non-native sound and recognizing it as being similar to a sound in their phonological category. This is an example of the first type of assimilation: clearly assimilated to a particular native segmental category. For example, although some of the phonemes for the stimulus [tʃʌɹʔtʃ] did not exist in the children's native dialect, they assimilated the sounds to the closest match and correctly recognized the word as *church.*

Also, when the children were unable to produce any response at all, it was most likely due to the sound being assimilated as an uncategorizable speech rather than a nonspeech sound. The sound was probably assimilated within the child's native phonological space as a "speechlike gestural constellation," but it did not fall into any clear category, which is why the child was unable to produce even an answer. It is highly doubtful that the child did not recognize the sounds he/she heard as sounds of human speech.

Although the younger children in the Nathan, Wells, and Donlan (1998) study did not understand the words pronounced in the non-native dialect as well as the older children, we know from the Chambers (1988) study that the children will eventually learn to associate meaning with the words in the non-native dialect. So although older children are better at the initial perception and understanding of words in a non-native dialect,

younger children can learn to understand the words with further exposure to the dialect. This exposure could lead to the features of the non-native dialect eventually becoming part of the child's native dialect.

ADULT DATA

Adult Study 1

The driving force for Labov's 1989 experiments on dialects was his observation that sound change was progressing rapidly in the major cities of the United States. Two of the major sound changes are the Northern Cities Shift, realized in Chicago, and the Southern Cities Shift realized in Birmingham, Alabama (Labov 1989). In light of these sound changes, Labov asked the question, "How do people from Birmingham understand Chicagoans?"(Labov 1989: 176) The hypotheses were:

1. They may have already built a pan-dialectal phonology that includes the Chicago realization of English vowels; *or*
2. they may deduce the systems by observing several correlated changes; *or*
3. failing to decode the vowels in an appropriate way, they may discard the vowel information and use morphological, syntactic, semantic and pragmatic information to deduce the meaning (Labov 1989: 176).

Labov expected to find that the third hypothesis, called parallel strategy, would prove to be the most prevalent method.

Labov mentions a pilot study he conducted back in 1985 using a technique called extended decoding. In this experiment, subjects heard a recording of spontaneous speech by a 13 year old Chicago girl. The listeners, from Philadelphia and Chicago, were asked

to repeat what they heard in the narrative. Labov found that the subjects who had been raised in the Northern Cities understood what was said in the narrative much better than those who hadn't been raised there. Labov also noticed that the parallel strategy that he had thought would be used, was not utilized by many subjects. The subjects didn't seem to take into account the context of the narrative to "overcome the effects of an unexpected phone" (Labov 1989: 179). Instead, subjects gave nonsense words as their interpretations. For example, the word *bedroom* was heard by subjects as *budgemen*, *budgement*, *budgepen*, and other non-words. Labov felt that "something in the stimulus had interfered with their ability to use the syntactic, semantic and pragmatic information that was available" (Labov 1989: 180).

The main research, done in 1989, consisted of four experiments: Extended Coding, Vowel Identification, Contextual Gating and Vowel Alteration. Although the other three tests provide useful information, we will focus on Labov's Gating Experiment. The experiment was structured by having subjects hear a list of 18 isolated words or syllables and after each word. They were then asked to write it out using regular spelling (non-IPA). Next, the subjects heard the same items in a phrase and were asked to transcribe what they had heard. Finally, the subjects were given a sheet with 18 sentences written out with blanks in the spots where the phrases they had just heard should have been. They then heard the full sentences, including the phrases, and were asked to fill in the blanks on the sheets that were given to them.

This gating experiment was done in Chicago, Birmingham, and Philadelphia. Philadelphia was part of the Middle Atlantic States that had their own distinctive dialect features along with features of the other two regions. There were 69 subjects from

Chicago, 121 from Birmingham, and 54 from Philadelphia who participated. Labov specifically looked at the phonemic contrasts of the fronting of short /o/ to short/æ/ and the backing of short /e/ to /ʌ/ in the Northern Cities dialect. Examples of what was given to the subjects are:

> Word: socks [sæks]
> Phrase: you had to wear socks [yɛdəwɚsæks]
> Sentence: You had to _____. No sandals. You had to [wɚrsæks]. No sandals.

For this particular example, the majority of subjects realized the word with an /æ/ phoneme and identified the word as *sacks*. The results become more interesting at the phrasal level. About one third of the Chicago subjects identified the word correctly as *socks*, while only two of the 15 Philadelphians and one of the 24 subjects from Birmingham identified it as such. When the full phrase with "No sandals" was added on, every Chicagoan except for two correctly identified *socks*. Of the subjects from Philadelphia and Birmingham, 30% to 50% still identified the wrong word, even though the sentence made no sense.

There were other words where the production of the vowels was similar in the Chicago and Birmingham regions, but completely different in Philadelphia. For those words, Chicagoans did the best at correctly identifying words; the subjects from Birmingham did almost as well; and the Philadelphians did the worst. The results for the rest of the 18 words were similar with the Chicagoans being able to correctly identify the words the best. The non-Chicagoans repeatedly identified the wrong word despite the addition of contextual information. There was only one case where almost everybody identified the correct word, and that word was *decker*. The word was used in the context

of the phrase *double decker bus* and subjects had a strong lexical reaction to pair those two words together.

This study also proved that the parallel theory of using context to decipher words in non-native dialects was not adequate to describe how people understood dialects. Labov's interpretation of the results of both the pilot study and this gating study is that "an aberrant phonetic form may completely block access to other sources of information relevant to the interpretation of the sentence as a whole" (Labov 1989: 184). He continues saying that "when a segment is clearly identified as a member of a given phoneme…it appears to be difficult for many listeners to discard that information" (Labov 1989: 188). Further on in Labov's experimentation, he found that "phonetically unclear utterances are actually easier to decipher than clear ones that contain unexpected phonetic forms" (Labov 1989: 194).

Labov states that "it is phonology rather than phonetics that presents the major problem for comprehension" (Labov 1989: 194). This means that it is not the actual sounds that a person hears that makes it difficult to comprehend a word, but rather the incorrect phonological categorizing of the sounds that proves to be the problem. He concludes by saying that "much of language processing occurs in a step-by-step deterministic manner, and that the output of one component of speech recognition process is not easily overridden by another" (Labov 1989: 196-197).

Adult Study 2

Labov, Karen and Miller (1991) conducted another set of experiments with the Philadelphia area dialect. In this dialect, there exists a near to complete merger of the /er/-

/ʌr/ contrast among most Philadelphians. This leads to a possible confusion of minimal pairs such as *ferry* and *furry*, *merry* and *Murray*, *kerry* and *curry*, and others of a similar sort.

In one part of an experiment, Labov, Karen and Miller constructed a Semantic Disambiguation test to check for whether or not listeners could hear this type of distinction. As presented in Labov, Karen Miller (1991: 58), it entailed:

1. Subjects are asked to listen to a narrative and give their own judgment as to what was the right or wrong thing to do, or whether the act performed was right or wrong.
2. In one sentence near the end of the narrative, a sentence is constructed that involves a crucial distinction being studied, so that if the one form or the other is used, the interpretation of the actual events being described will be radically different. Subjects are randomly given versions of the narrative with one or the other of the forms being tested.
3. The narrative is continued to the end with sentences that are all ambiguous in regard to the critical event and are necessarily interpreted by the listener in a way consistent with the interpretation made in (2).
4. Subjects are asked for their opinions, and the discussion is continued until it is clear which interpretation has been made in (2).

The story used was about a coach's decision and so the test was appropriately called "The Coach Test." The minimal pair used in this test was *Murray in* and *Merion*. The narrative included a boy and a girl who played for a local school baseball team. The boy, named Murray, was a bad player but was a hard worker. The girl, who was called

Merion because she was from Upper Merion Township, was a bad batter, but a good hitter. During an important game, the center fielder injured himself and the coach had the difficult decision of deciding who should take his spot. The narrative was recorded in a standard (non-Philadelphian) dialect with two possible choices:

A "I gotta play *Merion* there."

B "I gotta play *Murray in* there."

One ending was played for a subject and then the subject was then asked if the coach made the right decision. After the subject gave an answer, he/she was then played the last segment of the narrative with the alternate ending. For example, if the *Merion* ending was played the first time, then the *Murray in* ending was played the second time. It was understood that if the subjects had heard a difference between the two recordings, they would assume that their first interpretation was incorrect and then change their response to match the second ending when asked whether or not the coach had made the right decision.

The subjects listening to this narrative consisted of 21 students from the Philadelphia area and 15 from outside of the area. Of the 21 Philadelphians, 14 responded in a manner suggesting they heard the correct stimulus, while seven matched incorrectly. Of the non-Philadelphians, 14 were able to hear the distinction correctly while only one was unable to. This shows that the Philadelphians are worse that their non-Philadelphian peers in discriminating between /er/ and /ʌr/.

Another part of the experiment was a commutation test where the 36 subjects were recorded saying seven tokens of *ferry* and seven of *furry*. The subjects were then played back 10 of their words and were asked to identify whether what they heard was

the word *ferry* or *furry*. Of the 15 non-Philadelphians, all of them correctly labeled each word correctly. Of the 21 Philadelphians, they were further categorized to whether or not they themselves produced a near merger of /er/ and /ʌr/ in their own speech. 10 Philadelphians did not have the near merger, while the other 11 Philadelphians did. Of the 10 Philadelphians who did not the /er/ and /ʌr/ near merger, 79% of their answers to the commutation test were correctly matched. The other 11 Philadelphians who had the near merger only matched 51% of their own tokens correctly.

The general conclusions that Labov, Karen and Miller made from both parts of this experiment areas are as follows:

> Philadelphians with a full merger show the expected random response to self-commutation tests, and a severe reduction in the ability to categorize tokens that are clearly distinct.
>
> Philadelphians with a near-merger, either overlapping or nonoverlapping, do not show a significant improvement over speakers with a merger in their categorizations.
>
> Philadelphians with a clear distinction are significantly better than others in categorizing their own productions…
>
> All Philadelphians are worse than non-Philadelphians in judging their own productions or a standard clear distinction (1991:71).

Something Labov, Karen and Miller failed to report on in the Coach Test was how many of the responses were changed from the original when the alternate recording was played. Although one would assume that the ones who answered correctly had

changed from an incorrect to a correct choice upon hearing the different recording, what is to say that something different didn't happen? It could be that the subjects had originally heard the first recording incorrectly and responded with the wrong answer. Therefore, if they did not change the answer when the second recording was played, the original answer that they remained with would turn out to be right in the end.

ANALYSIS OF ADULT DATA

The data from Labov's (1989) experiment shows that when presented with a word pronounced in a non-native dialect, adults have a difficult time understanding and perceiving the correct meaning. Even in the presence of contextual information, there were still a large percentage of subjects continued to respond with lexical answers that made no sense in context. Although the presence of contextual information in this experiment yielded better results than if the study had been done without the contextual information, the fact remains that adults do a poor job in perception of non-native dialects.

These results appear to support Best's perceptual assimilation model in the methodology that was employed to identify the words. When the subjects were presented with words in a non-native dialect, they proceeded to assimilate them to the closest native category of phonemes in order to decide what the words were. Once the nearest match was found, the subjects were very certain that what they had heard and identified the word to be was in fact the correct word. When context was added, even if it conflicted with the word the subjects had identified or made no sense semantically, many of the subjects stood by their original identification.

We see that the assimilation to a phonological category takes precedence over pragmatic categorization. So as opposed to Labov's parallel strategy of "discard[ing] the vowel information and us[ing] morphological, syntactic, semantic and pragmatic information to deduce the meaning" of a word, subjects were more likely to discard the latter information from context than the phonetic information.

According to Labov (1989), once a word is identified as a member of a given phoneme, it is difficult to change one's categorization of the word. Therefore, in regards to Labov's (1991) second experiment with Karen and Miller, it would not be implausible to assume that a subject in this study could have identified one word when the first ending of *Merion/Murry in* was played and then upon hearing the second ending, responded with the same response as before. Although, it is important to note that the phonological stimulus was changed in this experiment unlike the previous study. Perhaps if the stimulus is changed, even if it is only a slight change, then reorganization of the original category is possible.

ISSUES IN COMPARING CHILDREN AND ADULTS

In comparing the children and adult studies, its interesting to note that while the children made many phonetic responses of just imitating back the words they had heard, there were no cases of the adults repeating back the word in the non-native dialect. This could be because adults have a larger inventory of lexical items. The children may have "defaulted" to phonetic responses when they could not identify similar words incorporating the phonemes in their phonological categorization. Adults on the other

hand have more combinations of phonemes in the form of words to possibly associate with the words in the non-native sounds.

However, if it is the case that the children gave a phonetic response despite the fact that they knew the correct meaning of the word, this could support for the idea raised before about volition and the preference. Children could be more open and willing to imitate a different dialect in order to "belong" and adapt to change. On the other hand, adults could be less willing to make changes in their dialect.

There have been other studies done to show that adults can "acquire" new dialects such as one by Munro, Derwing and Flege (1999). This specific study dealt with Canadian adults who had moved to Birmingham, Alabama. The results showed that there was some degree of dialect pronunciation that was acquired by the adults, but not in the phonological categorization. It is impossible for adults to learn to discriminate new phonemic contrasts.

The issue of how context plays a role in helping to deduce meaning from non-native dialect brings about an interesting question: if contextual information had been presented initially, what method would they have utilized to identify the words? Which would be the first employed, Best's perceptual assimilation model or Labov's parallel strategy. Best's model takes into account solely the phonemes of the words while Labov's model says that contextual information is the means by which words are correctly identified. It would be interesting to see that when given the choice, whether context would take precedence over the phonetic cues.

CONCLUSION

Comparing the studies on the perception of non-native dialects among children and adults, we can conclude that adults and older children have the advantage in identifying words spoken in a non-native dialect, but not in assimilating the new speech sounds into their own phonological store. Older subjects work only with their existing phonological categories to find a match for non-native speech sounds, but because there are more words in their lexical store, there is a possibility for more combinations of phonemes than a younger child may have. Also, adults can potentially deduce the correct meaning of a word based on contextual information, but we have seen how even context can fail. As disheartening as it may be, the fact remains that adults are limited in their capabilities to assimilate new non-native dialects.

While young children may have difficulty in their initial understanding of words spoken in a non-native dialect, they have the potential to learn. Through further exposure and experiences with various dialects, children can add the phonemic contrasts and features of those dialects to their own native categorization. Also, as a child learns more words, his/her ability to identify words in non-native dialects increases. Children also have an advantage in the capability to produce the sounds of non-native dialect and thereby assimilating the dialect into their own.

Although the degrees of perception of non-native dialects differ, children and adults appear to use the same method to understand what is being said. Best's perceptual assimilation model of matching similar existing phonological categories with the non-native sounds sufficiently describes how people understand dialect.

Bibliography

Best, C. T. (1994). The emergence of native-language phonological influences in infants: a perceptual assimilation model. In J. C. Goodman & H. C. Nusbaum (eds), *The development of speech perception*. Cambridge, Mass: MIT Press.

Best, C. T. (1995). A direct realist view of cross-language perception. In W. Strange (ed), *Speech perception and linguistic experience*. Timonium, MD: York Press.

Chambers, J. K. (1988). Acquisition of phonological variants. In A. R. Thomas (ed), *Methods in dialectology*. Clevedon, England: Multilingual Matters.

Chambers, J. K. (1992). Dialect acquisition. *Language*, 68, 4, 673-705.

Chambers, J. K. & Trudgill, P. (1980). *Dialectology*. Cambridge: Cambridge University Press.

Munro, M. J., Derwing, T. M., & Flege, J. E. (1999) Canadians in Alabama: a perceptual study of dialect acquisition in adults. *Journal of Phonetics*, 27, 385-403.

Labov, W. (1989). The limitations of context: evidence from misunderstandings in Chicago. *Chicago Linguistic Society*, 25, 171-200.

Labov, W., Karen, M., & Miller, C. (1991). Near-mergers and the suspension of phonemic contrast. *Language, Variation and Change*, 3, 33-74.

Nathan, L., Wells, B., & Donlan, C. (1998). Children's comprehension of unfamiliar regional accents: a preliminary investigation. *Journal of Child Language*, 25, 343-365.

Pisoni, D. B., Lively, S. E., & Logan, J. S. (1994) Perceptual learning of nonnative speech contrasts: implications for theories of speech perception. In J. C. Goodman & H. C. Nusbaum (eds), *The development of speech perception*. Cambridge, Mass: MIT Press.

Polka, L., Jusczyk, P. W., & Rvachew, S. (1995) Methods for studying speech perception in infants and children. In W. Strange (ed), *Speech perception and linguistic experience*. Baltimore, MD: York Press.

Wells, J. C. (1982). *Accents of English Volume 1: an introduction*. Cambridge: Cambridge University Press.

Werker, J. F. (1995). Age-related changes in cross-language speech perception: standing at the crossroads. In W. Strange (ed), *Speech perception and linguistic experience*. Timonium, MD: York Press.

Werker, J. F. & Pegg, J. E. (1992) Infant speech perception and phonological acquisition. In C. Ferguson, L. Menn, & C. Stoel-Gammon (eds), *Phonological development: models, research, and implications*. Timonium, MD: York Press.

Werker, J. F. & Tees, R. C. (1983). Developmental changes across childhood in the perception of non-native speech sounds. *Canadian Journal of Psychology*, 21, 83-101.

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