The Role of Segmental Information in Language Discrimination by English-learning 5-month-olds

Elizabeth K. Johnson, Peter W. Jusczyk, and Franck Ramus*

Introduction

In today's world, being raised in a monolingual environment is the exception rather than the rule. If infants are unable to discriminate between utterances belonging to different languages, growing up in a bilingual setting could seriously hinder language acquisition. Nonetheless, children routinely succeed at acquiring more than one language at once. Thus, it appears that infants are capable of separating input utterances according to their source language.

Research has suggested that language learners initially categorize languages according to their rhythmic characteristics. Different languages belong to different rhythmic families. For example, languages such as English and Dutch are described as stress-timed languages, whereas languages such as Japanese and Tamil are described as a mora-timed language. Romance languages such as Italian and Spanish fall into a third category described as syllable-timed. Mehler, Jusczyk, Lambertz, et al. (1998) provided evidence for the hypothesis that learners characterize languages based of rhythmic criteria by showing that French newborns can discriminate their native language from a language belonging to a different rhythm class, even if the language samples are filtered (low pass, 400 Hz) to remove most of the phonotactic information that could be used to distinguish the languages. Subsequent work by Nazzi, Bertoncini, and Mehler (1998) provided additional evidence for this hypothesis by showing that French newborns can discriminate between low-pass filtered recordings from two foreign languages belonging to different rhythmic classes (Japanese and English), but not two foreign languages belonging to the same rhythmic class (English and Dutch).

These results support the hypothesis that infants begin discriminating languages by paying attention to rhythmic information. However, as Nazzi et al. (1998) pointed out, these results do not explain when and how infants begin discriminating between languages belonging to the *same* rhythmic class. The finding that English-learning 2-month-olds do not discriminate between English and Dutch supports the view that discriminating between two languages belonging to the same rhythm class is difficult for infants (Christophe & Morton, 1998). Nonetheless, recent studies have shown that 4-month-olds in Barcelona can discriminate between two syllable-timed languages: Spanish and Catalan (Bosch, 1998). In addition, Nazzi, Jusczyk, and Johnson (2000) found that English-learning 5-month-olds are able to discriminate between two stress-timed languages: Dutch and English.

One possible explanation for these results is that as infants mature, their ability to discriminate all language pairs improves. This general improvement in the ability to discriminate languages would facilitate infants' ability to discriminate all language pairs,

---

* These data were compiled and written by Elizabeth K. Johnson (zab@jhu.edu) following the death of PWJ. The Saltanaj experiments were carried out by Elizabeth Johnson and Peter Jusczyk in collaboration with Franck Ramus. The low-pass filtering experiments were carried out by Peter Jusczyk and Elizabeth Johnson.

including pairs that belong to the same rhythm class. However, evidence provided by Nazzi et al. (2000) does not support this hypothesis. Although English-learning 5-month-olds can discriminate between their native language and another language in the same rhythm class (e.g. Dutch and English), they cannot discriminate between two unfamiliar languages belonging to a rhythm class different than that of their native language (e.g. Itallian and Spanish). This finding suggests the possibility that infants can only discriminate between two languages in the same rhythm class if both languages belong to the native rhythm class. However, the report that English-learning infants do not discriminate between two foreign languages from the native rhythm class (Dutch and German) provides evidence against this hypothesis (Nazzi et al., 2000). Therefore, Nazzi et al. suggested that although infants appear to initially use broad rhythmic class information to discriminate languages belonging to different rhythmic classes, they eventually learn to discriminate between languages within the native rhythm class by using language-specific knowledge about the detailed sound organization of their native language. Nazzi et al.'s finding that American English-learning infants discriminate between two dialects of English (British English and American English) support this view. In short, it appears that newborns can only discriminate languages with very different rhythmic characteristics. As infants gain more experience about with their native language, they learn enough about their native language that they can discriminate their native language from all other languages.

Determining what language-specific knowledge infants use to discriminate between their native language and another language belonging to the native rhythmic class is a problem that remains to be solved. Nazzi et al. (2000) argued that infants use language specific rhythmic information to discriminate American English from both Dutch and British English. One way to test this hypothesis would be to remove all segmental information from the passages used by Nazzi et al., and see if infants can still discriminate between Dutch and English. Traditionally, researchers have used low-pass filtering to try to remove segmental information from speech. For example, Mehler, Jusczyk, Lambertz, et al. (1988) and Nazzi et al. (1998) bolstered their claims that newborns use rhythmic information to discriminate between languages belonging to different rhythmic classes by showing that newborns discriminate between low-pass filtered recordings as readily as they discriminate between unfiltered recordings. But others have argued that low-pass filtering is not the best way to remove segmental information from speech (Ramus & Mehler, 1999). First, removing all energy above 400 Hz does not necessarily remove all segmental information from the speech signal. Second, rather than removing prosody, low-pass filtering often distorts it. Moreover, all tests of infants' language discrimination capabilities have used female voices, and both of these effects are exacerbated in female voices. Ramus and Mehler (1999) developed an alternative to low-pass filtering. This resynthesis technique is referred to as saltanaj. Natural speech is transformed into saltanaj speech by replacing all fricatives with the phoneme /s/, all stop consonants with /t/, all liquids with /l/, all glides with /j/, and all nasals with /n/. In addition, all vowels are replaced with /a/. Thus, this resynthesis technique removes phonetic and phonotactic information while leaving global intonation and syllabic rhythm intact.

Experiment 1

In the introduction, we reviewed evidence that English-learning 5-month-olds can discriminate between their native language and another language from the native rhythm class (e.g. English and Dutch). Nazzi et al. (2000) argued that infants achieve this by using language-specific knowledge about the sound structure of English. In particular, Nazzi et al. suggest that infants use language-specific prosodic or rhythmic, rather than segmental, information to discriminate between Dutch and English. In Experiment 1, we test the hypothesis that English-learning infants can discriminate English and Dutch even when the segmental cues to language identity have been removed. Since low-pass filtering may not be the best way to remove segmental information from the speech signal, we used resynthesized versions of the English and Dutch passages used by Nazzie et al. (2000). These passages were resythesized using the saltanaj technique described by Ramus and Mehler (1999). If English-learning 5-month-olds do not rely on segmental cues to discriminate English and Dutch, then they should be able to discriminate saltanaj Dutch from saltanaj English (infants familiarized with Dutch should prefer to listen to English during the test phase, and vice versa). If infants fail to discriminate saltanaj Dutch from saltanaj English, there are at least two possible explanations. Either English-learning 5-month-olds use segmental information to discriminate English and Dutch, or infants do not process saltanaj in the same way that they process natural speech.

Methods

Participants

Twenty English-learning 5-month-olds from the Baltimore-Annapolis region were tested (13 Males and 7 females). The infants were approximately 5.0 months old (range 4.5-5.5), with a mean age of 153 days (Range: 139-167); SD=9.5). The data from 9 additional infants were excluded for the following reasons: fussiness (7), would not turn to lights (1), looking times less than three seconds (1). Participants were recruited to come to the lab via letters mailed by the Maryland State Health Department, advertisements in local parenting magazines, and referrals from past participants. Parental consent was obtained for all participants. In appreciation for their participation, all infants were given a small toy or certificate and a $10 travel reimbursement.

Stimuli

The 16 passages used in this experiment were synthesized versions of the 8 Dutch and 8 British English passages used in Nazzi et al (1998; 2000). The 8 passages were produced by four different female speakers of two languages (two passages per speaker). Speakers were chosen so as to minimize voice quality differences (pitch and timbre) between speakers. All passages were approximately 16 seconds long, and were closely matched in number of syllables per sentence. Each passage contained 5 unrelated sentences, and was recorded in adult-directed speech. The saltanaj versions of these passages were synthesized using the procedure described by Ramus and Mehler (1999).

Design

Ten infants were randomly assigned to each of two groups, defined in terms of the language presented during the familiarization (British English or Dutch).

Procedure and Apparatus

Infants were tested using the same version of the Headturn Preference procedure used by Nazzi et al. (2000). In this procedure, infants sit in the center of a caregiver's lap. The caregiver is seated on a chair in the center of a three-sided booth made of white pegboard. A red light and a speaker are mounted at eye level on the center of each side panel, and a green light is located at eye level on the center of the front panel. The green light flashes at the start of each trial. Once the infant orients towards the green light, the green light stops flashing and one of the two red side lights immediately begins flashing. Once the infant orients towards the flashing light, a soundfile begins playing from the speaker mounted directly behind the light. The soundfile will continue to play until the infant looks away for more than two seconds, or the soundfile ends. Thus, the infant essentially controls how long he or she hears the soundfiles. Differential preference for one type of soundfile over another is typically used as evidence that infants can detect a difference between the types of soundfiles.

As in Nazzi et al. (2000), this experiment consisted of both a familiarization and test phase. During the familiarization, four passages in one language (two different passages from two different speakers of either Dutch or English) were presented until infants accrued 20 seconds of orientation time to each passage. During the test phase, infants heard four passages from the language they were not familiarized with (spoken by two speakers) as well as four new passages by two new speakers in the familiarized language. The order of soundfile-presentation during both phases was randomized for each infant. The side of the booth from which the stimuli was played was randomized for each infant too. The experimenter and a video camera were hidden behind the front panel of the booth. The experimenter recorded the direction and duration of infants' orientation via a botton box connected to a Macintosh computer. Computer software was responsible for the selection and randomization of the stimuli and for the termination of test trials. Both the caregiver and the experimenter wore tight-fitting headphones over which loud masking music was played (for a more thorough methodological discussion, see Kemler Nelson et al., 1995). As in Nazzi et al. (2000) as well as Bosch & Sebastian-Galles (1997; Bosch, 1998), we expected that if infants could discriminate between the languages, then they would listen longer to the passages from the new (unfamilarized) language during the test phase.

All recordings were played at a comfortable listening level (approximately 72 dB SPL, according to a Quest (Model 215) sound meter. The audio output was generated from the digitized waveforms of the samples. A 16-bit D/A converter was used to recreate the audio signal. The output was fed through anti-aliasing filters and a Kenwood audio amplifier (KA 5700) to the two 7-inch Cambridge Soundworks loudspeakers mounted on the side walls of the testing booth.

Results and Discussion

Mean orientation times to the two languages during the test phase were calculated for each infant. Twelve out of 20 infants had longer orientation times for the passages in the nonfamiliarized language. Across all infants, the average orientation times were 6.4 seconds (SD = 1.8) for the passages in the new language and 6.13 (SD = 2.8) for the passages in the familiarized language. A 2 X 2 mixed design ANOVA with test item (familiarized language versus nonfamiliarized language) and familiarization condition (Dutch versus English) revealed no significant effect of test item ($F<1$). In addition, there was no main effect of familiarization condition ($F < 1$), or interaction between test item and familiarization condition ($F < 1$). Thus, it appears that infants in this experiment did not discriminate between resynthesized Dutch and English: two languages from their native rhythm class. As mentioned above, there are at least two possible explanations for this result. Either English-learning 5-month-olds rely on segmental information to discriminate between English and Dutch, or infants do not process saltanaj speech in the same way that they process natural speech.

Experiment 2

In Experiment 1, we used the saltanaj resynthesis method to try to remove the segmental information from the Dutch and English passages used by Nazzi et al. (2000). Infants tested on these resynthesized passages failed to discriminate between English and Dutch. In Experiment 2, we use low-pass filtering rather than resynthesis to remove most of the segmental information from the same English and Dutch passages used in Experiment 1. If infants once again fail to discriminate between English and Dutch, then this would lend support to the hypothesis that English-learning infants need segmental information to discriminate between English and Dutch.

Methods

Subjects

Twenty English-learning 5-month-olds from the Baltimore-Annapolis region were tested (12 Males and 8 females). The infants were approximately 5.0 months old (range 4.5-5.5), with a mean age of 154 days (Range: 138-166); SD=8.7). The data from 12 additional infants were excluded for the following reasons: fussiness (7), falling asleep (1), bilingual home environment (1), looking times less than three seconds (3). Participants were recruited to come to the lab via letters mailed by the Maryland State Health Department, advertisements in local parenting magazines, and referrals from past participants. Parental consent was obtained for all participants. In appreciation for their participation, all infants were given a small toy or certificate and a $10 travel reimbursement.

Stimuli

The same passages were used as in Experiment 1. However, after the conversion of the digitized files to audio output via the D/A converter, the audio signals were passed through a Krohn-Hite filter with the low-pass cutoff set at 400 Hz with an attenuation slope of 48 dB per octave. This filter level eliminated most phonetic information while leaving the intonation, stress, and rhythm of the passages fairly intact. The filtered soundfiles were output to the amplifier and then to the loudspeakers in the testing room. Since low-pass filtering reduces the overall amplitude of speech sounds, the loudness levels on the amplifiers were adjusted to ensure that the passages were played at the same volume as those used in Experiment 1 (approximately 72 dB SPL, as measured by a Quest Model 215 sound meter).

Design

Ten infants were randomly assigned to each of two groups, defined in terms of the language presented during the familiarization (British English or Dutch).

Procedure and Apparatus

The same procedure and apparatus were used as in Experiment 1.

Results and Discussion

Mean orientation times to the two languages during the test phase were calculated for each infant. Nine out of 20 infants had longer orientation times for the passages in the nonfamiliarized language. Across all infants, the average orientation times were 7.05 seconds (SD = 3.2) for the passages in the new language and 7.05 (SD = 3.1) for the passages in the familiarized language. A 2 X 2 mixed design ANOVA with test item (familiarized language versus nonfamiliarized language) and familiarization condition (Dutch versus English) revealed no significant effect of test item ($F<1$). In addition, there was no main effect of familiarization condition ($F < 1$), or interaction between test item and familiarization condition ($F < 1$). Thus, it appears that infants in this experiment could not discriminate between low-pass filtered Dutch and English: two languages from their native rhythm class. This finding lends support to the hypothesis that English-learning 5-month-olds need segmental information to discriminate English and Dutch.

Experiment 3

The results of Experiments 1 and 2 suggest that English-learning 5-month-olds use segmental information to discriminate English and Dutch. However, before taking these two null results as strong support for 5-month-olds' use of segmental information, it is important to make sure that 5-month-olds have any ability whatsoever to discriminate resynthesized or low-pass filtered language samples. Therefore, in Experiment 3, we will test infants' ability to discriminate saltanaj English from saltanaj Japanese. These two languages differ greatly in their rhythmic and prosodic structure, and should therefore be easy to discriminate in the absence of segmental information. If English-learning 5-month-olds discriminate saltanaj English from saltanaj Japanese, then this finding would support the hypothesis that infants' failure to discriminate English and Dutch in Experiments 1 and 2 was due to the removal of segmental information from the passages rather than some artificat caused by the use of resynthesized or low-pass filtered speech.

Methods

Subjects

Twenty English-learning 5-month-olds from the Baltimore-Annapolis region were tested (11 Males and 9 females). The infants were approximately 5.0 months old (range 4.5-5.5), with a mean age of 152 days (Range: 136-168); SD=10.4). The data from 11 additional infants were excluded due to fussiness. Participants were recruited to come to the lab via letters mailed by the Maryland State Health Department, advertisements in local parenting magazines, and referrals from past participants. Parental consent was obtained for all participants. In appreciation for their participation, all infants were given a small toy or certificate and a $10 travel reimbursement.

Stimuli

The 16 passages used in this experiment were synthesized versions of the 8 Japanese and 8 British English passages used in Nazzi et al (2000). Speakers were chosen so as to minimize voice quality differences (pitch and timbre) between speakers. All passages were approximately 16 seconds long, and were closely matched in number of syllables. Each passage contained 5 unrelated sentences, and was recorded in adult-directed speech. The saltanaj versions of these passages were synthesized using the procedure described by Ramus and Mehler (1999).

Procedure, Apparatus, and Design

The same procedure and apparatus were used as in Experiment 1. Ten infants were randomly assigned to each of two groups, defined in terms of the language presented during the familiarization (British English or Japanese).

Results and Discussion

Mean orientation times to the two languages during the test phase were calculated for each infant. Eight out of 20 infants had longer orientation times for the passages in the nonfamiliarized language. Across all infants, the average orientation times were 7.17 seconds (SD = 2.94) for the passages in the new language and 7.6 (SD = 2.7) for the passages in the familiarized language. A 2 X 2 mixed design ANOVA with test item (familiarized language versus nonfamiliarized language) and familiarization condition (Japanese versus English) revealed no significant effect of test item ($F<1$). In addition, there was no main effect of familiarization condition ($F < 1$), or interaction between test item and familiarization condition ($F < 1$). Thus, it appears that infants in this experiment could not discriminate between resynthesized Japanese and English: two languages from different rhythm classes. Given past findings showing that French newborns can discriminate low-pass filtered Japanese and English (Nazzi, 1998), this result forces us to question whether 5-month-olds process saltanaj speech samples in the same way that they process natural speech.

Experiment 4

The results of Experiment 3 suggest that infants do not necessarily process resynthesized speech in the same way they process untransformed speech. In the following study, we test whether infants discriminate between samples of low-pass filtered speech more readily than saltanaj speech. Past studies using the High Amplitude Sucking Paradigm have shown that French newborns can discriminate low-pass filtered Japanese and English (Nazzi et a., 1998). Therefore, it would be surprising to see that English-learning 5-month-olds cannot discriminate low-pass filtered speech. If infants in Experiment 4 succeed at discriminating low-pass filtered Japanese and English, then this would indicate that segmental information may be more important for discriminating two languages from the same rhythm class than two languages from different rhythm classes (recall that infants in Experiment 2 failed to segment low-pass filtered Dutch and English). On the other hand, if infants fail to discriminate low-pass filtered Japanese and English, then this might suggest that the design of the experiments included in this study is not appropriate for testing language discrimination capabilities in 5-month-olds.

Methods

Subjects

Twenty English-learning 5-month-olds from the Baltimore-Annapolis region were tested (11 Males and 9 females). The infants were approximately 5.0 months old (range 4.5-5.5), with a mean age of 147 days (Range: 122-170; SD=9.3). The data from 8 additional infants were excluded for the following reasons: fussiness (5), failure to look towards the side lights (2), and parental interference (1). Participants were recruited to come to the lab via letters mailed by the Maryland State Health Department, advertisements in local parenting magazines, and referrals from past participants. Parental consent was obtained for all participants. In appreciation for their participation, all infants were given a small toy or certificate and a $10 travel reimbursement.

Stimuli

This experiment used the same passages used in Experiment 3. However, instead of resynthesizing the passages, the original passages were low-pass filtered using the same method described in Experiment 2.

Design

Ten infants were randomly assigned to each of two groups, defined in terms of the language presented during the familiarization (British English or Japanese).

Procedure and Apparatus

The same procedure and apparatus were used as in Experiment 1.

Results and Discussion

Mean orientation times to the two languages during the test phase were calculated for each infant. Thirteen out of 20 infants had longer orientation times for the passages in the nonfamiliarized language. Across all infants, the average orientation times were 8.6

seconds (SD = 3.6) for the passages in the new language and 6.73 seconds (SD = 3.6) for the passages in the familiarized language. A 2 X 2 mixed design ANOVA with test item (familiarized language versus nonfamiliarized language) and familiarization condition (Japanese versus English) revealed a significant effect of test item, $F(1,18) = 7.3$, $p < .05$. There was no main effect of familiarization condition, $F(1, 18) = 2.5$, $p > .1$. However, there was a marginally significant interaction between test item and familiarization condition, $F(1, 18) = 3.4$, $p = .08$. This marginal effect was due to greater orientation differences in infants familiarized with Japanese (mean orientation time to nonfamiliarized language 3.52 seconds greater than that to the familiarized language) as opposed to English (mean orientation time to nonfamiliarized language only 0.07 seconds greater than that to the familiarized language). These results indicate that although infants discriminated between low-pass filtered Japanese and English, they may have done so better when they were familiarized with Japanese rather than English.

General Discussion

In Experiments 1 and 3, English-learning 5-month-olds failed to discriminate saltanaj English from either saltanaj Japanese or saltanaj Dutch. In Experiment 2, English-learning 5-month-olds failed to discriminate low-pass filtered Dutch from low-pass filtered English. However, in Experiment 4, infants did discriminate between low-pass filtered Japanese and English. In combination, these results do not provide any support to Nazzi et al.'s hypothesis that English-learning infants use prosodic or rhythmic information to discriminate Dutch and English. In addition, these results suggest that saltanaj speech samples do not carry the same information as low-pass filtered speech. Infants' discriminated between Japanese and English only when low-pass filtered speech was used. It is possible that low-pass filtering left behind some segmental cues to language identity, and infants used this segmental information to discriminate between Japanese and English. However, if this was true, it is hard to understand why infants did not use the same method to discriminate between Dutch and English in Experiment 2. Another possibility is that saltanaj sounded so harsh and unnatural to infants that they failed to process it as speech.

Another potentially interesting aspect of Experiment 4 is the marginal interaction between test item and familiarization condition. Infants had longer looking time differences if they were familiarized with Japanese rather than English. One possible explanation for this finding is that infants naturally have a preference for their own native language. Thus, when we familiarized English-learning infants with Japanese and then tested them on Japanese and English, there listening preferences could have been influenced by two converging factors: both a preference to listen to their own native language as well as a novelty effect driven by their familiarization with Japanese. On the other hand, when infants were familiarized with English, we pitted infants' preference to listen to their own native language against a novelty effect driven by their familiarization with English. If this was the case, then this would explain why the looking time differences were not as great when infants were familiarized with English rather than Japanese. It is interesting to note that in both low-pass filtering experiments (Dutch versus English and Japanese versus English), there was a trend during the test phase for infants to listen longer to the English passages over the passages in the foreign language (collapsing across familiarization condition). This trend was non-existent in Experiments 1 and 3 (saltanaj English versus saltanaj Dutch and saltanaj English versus saltanaj Japanese). In fact, a 2 X 2 mixed design ANOVA, with experiment type (saltanaj versus low-pass filter) as a between subjects measure and test passage type (English versus foreign language) as a within subjects measure, revealed a significant interaction between experiment type and test passage type. There was no significant main effect of either experiment type or test passage type ($F < 1$). The looking times observed during the test phase of the two low-pass filter experiments were subjected to a one-way repeated measures ANOVA, revealing that this effect was driven by infants' overall preference for English passages during the test phase, $F (1, 39) = 4.3$, $p < .05$. A similar analysis carried out on the looking times observed during the test phase of the saltanaj experiments revealed no such effect ($F < 1$). Thus, it appears that infants only preferred to listen to English in the low-pass filter studies. Once again, this is further evidence infants processed the resynthesized speech differently than the filtered speech. In addition, since infants showed this preference in both low-pass filter studies (English versus Dutch as well as English versus Japanese), it is possible that if we'd used a more sensitive technique, we may have found evidence that 5-month-old English-learning infants can discriminate low-pass filtered English and Dutch. Given infants apparent preference to listen to their native language, it might be interesting to carry out a straight-forward preference study (with no familiarization) to see if English-learning infants prefer to listen to low-pass filtered English over low-pass filtered Dutch. If so, this would provide evidence that segmental information may not be necessary for 5-month-olds to discriminate Dutch from English.

References

Bosch, L. (1998). Bilingual exposure and some consequences on native language recognition processes at four months. In P.K. Kuhl & L. Crum (Eds.), Proceedings of the Joint Meeting of the Acoustical Society of America and the International Conference on Acoustics (pp. 1599-1600). Seattle: Acoustical Society of America.

Bosch, L. & Sebastian-Galles, N. (1997). Native-language recognition abilities in 4-month-old infants from monolingual and bilingual environments. Cognition, 65, 33-69.

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

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.

Nazzi, T., Jusczyk, P.W., & Johnson, E.K. (2000). Language Discrimination by English-learning 5-month-olds: Effect of rhythm and familiarity. Journal of Memory and Language, 43, 1-19.

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.

Ramus, F., & Mehler, J. (1999). Language identification with suprasegmental cues: A study based on speech resynthesis. Journal of the Acoustical Society of America, 105 (1), 512-521.