Birds and babies : a comparison of the early development in vocal learners

Haar, S.M. ter

Citation

Haar, S. M. ter. (2013, June 5). *Birds and babies : a comparison of the early development in vocal learners*. Retrieved from https://hdl.handle.net/1887/20932

Author: Haar, Sita Minke ter
Title: Birds and babies : a comparison of the early development in vocal learners
Issue Date: 2013-06-05

5 Summary and General discussion

5.1 Summary

The aim of this thesis is to disentangle experience dependent and –independent mechanisms in vocal learning and to give insight in their relation to typological patterns in language and birdsong. In perceptual experiments on human infants and juvenile zebra finches I tested to which sounds learners attend and what drives this attention.

In Chapter 2, I tested human infants for their preferences independent of exposure by balancing stimuli for frequency of occurrence in the ambient language (in this case Dutch). The results show that 9-month-olds prefer homorganic syllables to heterorganic ones independent of frequency of occurrence. This probably reflects a processing of larger units (words or syllables) than segments, as has been suggested for early child language production. An early suprasegmental representation would also explain the lack of preference for unmarked segments at 9 months of age. The preference for heterorganic syllables in 12-month-olds possibly indicates a perceptual shift from processing larger units to segmental processing. Whether this is driven by biological maturation or exposure to language is currently unclear, however a combination of both factors is likely. Homorganic consonant vowel sequences have been suggested to be more common across languages, which could be partly caused by the early bias. The evidence for this cross-linguistic pattern in the literature is mixed, however, thus more insight in this is necessary to reveal if and how the homorganicity bias affects cross-linguistic patterns. Initial perceptual preferences in zebra finches (Chapter 3) were independent of song exposure because the birds were raised by the female parent only, which does not sing. These birds, naive to song, prefer common over uncommon song elements, somewhat similar to unmarked and marked sounds in human language. These preferences indicate that, independent of exposure, perception is selective for certain within species elements over other within-species elements.

In addition I tested experience dependent factors in both species. In human infants (Chapter 2) I tested whether frequent syllables are perceptually preferred over infrequent ones, independent of the possibly interfering factor markedness. Effects for frequency were found neither in 9-month-olds nor in 12-month-olds. This result contradicts earlier findings of infant preferences for frequent phonotactic patterns. However, the earlier study did not control for markedness, which demonstrates how important it is to study these effects separately. I do not deny that frequency or other forms of experience play a role, but I do want to emphasize the possibility of interaction with other factors such as markedness or homorganicity. More research is necessary to find out how exactly experience dependent and independent factors interact and at which stage in development.

Experience in zebra finches (Chapter 3) was manipulated by tape-tutoring the birds with songs containing different types of song elements (common or uncommon). Birds tutored with common songs preferred common songs at a later age as well, but now only preferred the specific common song they had been exposed to. Birds tutored with uncommon song changed their preference towards uncommon song and also specifically preferred their (uncommon) tutor song. This clearly demonstrates the effect of experience, which alters the preference towards the sounds they have been exposed to. The similarity between the pupil's song and the tutor song is higher for birds exposed to common songs than for birds exposed to uncommon songs, revealing an interaction between the early bias and later exposure. This interaction may explain why variation and uniformity in vocal signals exist.

The classification of elements into common and uncommon in the experiment with juvenile songbirds in Chapter 3 was based on literature describing zebra finch elements. These studies are valuable, but the methods used in the different studies are quite different and information on different populations is limited. In order to gain more insight in this, I studied the distribution of song elements and sequences across 13 zebra finch populations to approach a zebra finch typology (Chapter 4). Little evidence was shown for cross-population patterns of element sequences. Some populations however, showed weak evidence for some preferred sequences. Furthermore, one type of element seems more likely to be placed near or at the end of the song than at the beginning. Data on element distribution, however, did show that all element clusters occur in all populations. Nevertheless, variation between populations was found in the frequency of occurrence of certain elements within populations. A clear link with the experimental data in juvenile zebra finches could not be shown.

Further research should reveal what the long lasting effect of biases and experience is on variation and uniformity in birdsong and language. The two experimental studies do show a link between early biases and later production. The experiment on juvenile zebra finches shows how experience dependent and independent mechanisms interact. Furthermore both studies show that different factors affect perception at different developmental stages. The findings in this thesis, together with those in previous research, suggest that similar mechanisms operate in songbirds and human infants. More specifically, early biases are independent of experience and later preferences are more influenced by experience. Both factors seem to affect later production patterns, but the relation to cross-population and cross-linguistic patterns is not clear yet.

5.2 Experience independent processes in infants and birds

In Chapter 2 and 3 I demonstrate that both in human infants and in juvenile male songbirds, initial preferences are present independent of adult exposure. Using a 'looking while listening' paradigm, I show that human 9-month-old infants prefer homorganic syllables to heterorganic ones, independent of frequency. These syllables, of which the consonant (C) and the vowel (V) are pronounced with a similar place of articulation (PoA), are abundant in child language production. The early perceptual bias is interpreted as an indication of processing at a suprasegmental level, which has been suggested for child language production as well. Canonical babbling (repeated CV sequences like bababa) usually consists of homorganic syllables. In such sequences the syllable is called a 'frame' and only later the 'content' (segments) will develop (MacNeilage, 1998b). When young children start to produce their first words, these are also predominantly homorganic forms (Fikkert & Levelt, 2008). As shown in chapter 2, by 12 months of age (approximately the onset of word production) perceptual preference has shifted to heterorganic over homorganic syllables, which is interpreted as a first step in a developmental change from a suprasegmental to a segmental representation. The first heterorganic word productions have been found only around 17 months of age (Fikkert & Levelt, 2008) following the same chronology as perception but later. Thus the same developmental path is shown in perception and production, although perception precedes production. This would also be in line with the finding in chapter 2 that infants showed no sensitivity to segmental markedness contrasts at 9 months of age. If infants have no segmental representation yet, it is not likely that they will show a preference for segmental markedness. At 12 months of age however, the infants did not show a segmental markedness preference either, whereas their preference for heterorganic syllables suggest a segmental representation. A possible explanation is that a segmental representation only *begins* to form around 12 months of age but is not mature enough to allow for segmental markedness sensitivity. Further research is necessary to determine if infants perceptually prefer unmarked segments at a different age.

The finding of a perceptual preference independent of frequency suggests that phonological development is influenced by more than just language input. Chapter 3 demonstrates that also in male zebra finches early biases independent of song exposure are present. In this experiment the males were raised without song, which makes interpretation of the data slightly easier. A phonotaxis setup (not to be confused with the term phonotactics in the linguistic literature) was used in this experiment to measure preferences. A measure for preference is the time a bird spends on the side of the cage where a stimulus is played compared to the other side of the cage, where another stimulus is played. This indicated that birds naive to song, preferred 'typologically' common zebra finch elements to uncommon ones. Note that common and uncommon does not refer to how often the birds have heard the elements, because they haven't been exposed to song. Conceptually, the terms uncommon and common are more comparable to universal and non-universal, or unmarked and marked. Chapter 4 however, suggests that strictly speaking there is no evidence for non-universals. Nevertheless, looking at song described in the literature it seems that some elements are more likely to appear in songs (common) than others (uncommon). Thus before song exposure, male zebra finches attend more to common elements than uncommon ones.

For zebra finches we cannot completely exclude the role of exposure to vocalizations other than song. Females do not sing but they do produce calls, which may affect the perception of the nestlings. This possibility is discussed in the next section but we consider it unlikely that this would explain the early bias fully. For the following discussion I will therefore assume that this bias is a predisposition.

Birds tutored with common elements sing songs that are more accurately copied than those of birds tutored with uncommon elements. This shows again a link between the early bias and later production, similar to what is found for human infants. In previous literature a preference for conspecific over heterospecific songs was shown that was independent of song experience (Braaten & Reynolds, 1999; Nelson, 2000; Nelson & Marler, 1993; Soha & Marler, 2001) and this was hypothesized to be indicative of a selective learning mechanism. The perceptual system selectively directs attention to conspecific sounds, catalyzing song learning of the appropriate song elements. Marler et al suggest it may work as a template (see Adret 2004 of a review) of features for conspecific sounds pre-encoded in the brain. This results in selective neural activity triggered by auditory exposure corresponding to the conspecific features represented in the template (Marler & Peters, 1977; Marler & Sherman, 1983).

The findings in this thesis more specifically show that also *within species* selection is taking place. This also sheds light on the template hypothesis. Since not all within species elements are preferred initially, it is possible that the template does not cover features for the whole repertoire but only features for common elements and the rest is learnt. In other words, the more acoustic features match the template, the more it is activated and the stronger the bias. Thus, heterospecific elements match the template poorly whereas uncommon zebra finch elements match slightly better and common zebra finch elements show the best match and therefore trigger the most activation. This activation will enhance learning initially and subsequent exposure to non-matching elements will gradually adjust to or complement the initial template. This latter idea would also support the finding of a more gradual distribution of common and uncommon elements rather than a strict separation between the two. It is at this point unclear how the template is affected by experience and what is exactly encoded in the template. Nevertheless, the finding of an initial bias for within species vocalizations, suggest that not all conspecific vocalizations are pre-encoded equally.

A similar mechanism has been suggested for human infants as well. It has been suggested that (a sensitivity to) a set of unmarked phonological features is present in infants before they start the language acquisition process (Jakobson, 1941). This could be similar to the pre-encoded template for conspecific common elements in songbirds. However, we did not find a perceptual preference for unmarked sounds in infants in the presents study. Hence, at this point I cannot draw the conclusion that infants have a ‘template’ for unmarked features. Infants do have a bias for homorganicity (larger units than the segment) on the other hand (which has been suggested to be unmarked in itself) and this might be pre-encoded. A bias for processing units larger than the segment might interact with a bias for segmental phonological features and could possibly mask an effect at the ages tested in the experiments described in chapter 2. A preference for marked segments in older children at the stage of segmental processing, or a preference for homorganic syllables with unmarked PoA over homorganic syllables with a marked PoA could be a sign of a 'template' for unmarked features in infants. This has so far not been tested, but this would be interesting future research.

5.2.1 Alternative explanations for the early biases

The specific perceptual bias demonstrated in this thesis cannot be explained by a difference between native and foreign language since all stimuli were Dutch syllables. The bias cannot be explained by within language frequency either, since homorganic-heterorganic stimulus sets were balanced for frequency. However, self perception may play a role, since canonical babbling starts around 7 months of age and is thus quite common at 9 months of age. One preliminary finding (unpublished data from the same children tested in the experiment in Chapter 2) suggests that the babbling in 12-month-olds is more homorganic than in 9-month-olds, whereas the perceptual data show the opposite. This is in line with the idea that perception precedes production. The data on children's first words are also more in favor of the thought of production after perception, since heterorganic word productions are found first by 17 months of age (Fikkert & Levelt, 2008). More research is necessary however to unravel the exact interaction between perception and production.

For the initial bias for common elements found in zebra finches we cannot completely rule out the role of any type of exposure (Chapter 3). Females do not sing but they do produce calls (Zann 1996), which may affect the perception of the nestlings. Even though there is some similarity between female calls and the elements classified as common, not all of these elements resemble female calls. It is thus unlikely that female calls would explain the bias as such, but they may have an amplifying effect for instance. More importantly, for both human infants and juvenile zebra finches, we can conclude that the early biases cannot be explained by, respectively, adult language input and song input alone (Chapter 2 and 3 of this thesis). It is, thus, important to take other factors than language and song exposure into account when studying vocal acquisition.

5.3 Experience dependent processes in infants and birds

Previous research has shown that 9-month-old infants' discriminative abilities are affected by exposure to language (Best & McRoberts, 2003; Best et al, 1995; Kuhl et al, 2006; Werker & Tees, 1984) and that infants prefer to listen to their native language (Jusczyk et al, 1993a; Jusczyk et al, 1993b).

In Chapter 2 the role of experience is also specifically addressed by testing if infants prefer frequent over infrequent syllables independent of the possibly correlating factor markedness**.** Using the same paradigm as for experience independent factors, it was shown that 9-month-old infants do not prefer frequent over infrequent syllables. At 12 months of age, infants still do not prefer frequent over infrequent syllables. This is in contrast with earlier findings by Jusczyk et al., who demonstrated a perceptual preference for high probability phonotactic patterns within the native language in 9-month-olds. This can probably be explained by the selection of stimuli. The stimuli used by Jusczyk et al. were not balanced for markedness. First of all this resulted in a stronger frequency difference between frequent and infrequent stimulus sets, because it is easier to find highly frequent syllables or phonemes when it is no problem that they are unmarked. The extremely infrequent stimuli used by Jusczyk et al. might actually have been completely unfamiliar and perceptually comparable to foreign language to infants. Second, markedness itself might have played a role, interacting with frequency. That is, syllables that are both frequent AND unmarked could elicit a stronger response than frequent-marked or infrequent-unmarked syllables and the weakest response would be expected for infrequent-marked syllables. This means that the findings in the study by Jusczyk et al. might be confounded since not only frequency but also markedness may have caused the preference. Markedness alone, however, cannot explain the different results in the two studies since in this thesis the 9-month-olds showed no preference for unmarked segments (neither did the 12-month-olds). Further research is necessary to explore if and how markedness and frequency interact.

Experience does play a role in later perceptual development in zebra finches, as clearly indicated by the shift in perceptual preference in the group of zebra finches tutored with uncommon elements. These birds initially did not have a preference for uncommon elements, but after 20 days of exposure to uncommon elements, the initial preference for common elements declined and the birds preferred the song they had been exposed to. Also in the birds exposed to common elements, the preference became more specific for their tutor song: for common songs other than their tutor song, a preference was no longer found at 55 days post hatch. Several earlier studies have shown the relevance of auditory experience to song learning in birds (see Woolley 2012 for review) and have shown preferences for the tutor song specifically (Clayton, 1988; Houx & ten Cate, 1999; Houx & ten Cate, 1999b; Riebel & Smallegange, 2003; Riebel et al, 2002). Little is known however, on how predispositions and exposure interact. In chapter 3, I show that experience can override initial biases and change perceptual preferences. Nevertheless, the results on song production also indicate that even though production is affected by experience, there still remains an effect of the early bias on production as well. The perceptual preference is equally affected by exposure to common as well as uncommon elements, but the song quality (song similarity with the tutor) is better for birds tutored with common elements than for birds tutored with uncommon elements. Even though the lack of effect for frequency on infant preferences in Chapter 2 does not confirm experience dependent effects in human infants, obviously experience has to play a role at some point. Previous research also shows developmental changes from experience independent towards experience dependent effects. For instance, infants can discriminate non-native phonemes early in development whereas in the second half of the first year, this discriminative ability declines, while native language discrimination increases during this period. Whether the shift in preference from homorganic to heterorganic syllables (or from suprasegmental to segmental representations) between 9 and 12 months of age (Chapter 2) is driven by the same mechanism remains an open question. In human infants, it is often hard to distinguish between developmental change caused by auditory experience or by maturational processes. This is still an issue for the change in preference from homorganic to heterorganic syllables shown in chapter 2. Since the frequency for homorganic and heterorganic stimuli was balanced, this could not directly explain the shift. Obviously, 12-month-olds have 3 months more experience, which possibly gives them time to pass through the homorganic stage, and facilitates focusing on a new learning stage. Previous research comparing full term infants with preterm infants gives some insight in maturational processes, because their linguistic input is similar but the brain maturation stage is different (Pena et al, 2010). It would be interesting to test pre-term and full-term infants on their preferences for homorganicity.

In summary, combining knowledge from previous studies and the present research it becomes clear that in both human infants and zebra finches experience plays a role in vocal development and may partly override initial biases. It seems that initially in both species experience independent processes dominate, whereas over the course of development experience dependent processes take over. As is especially evident in the songbirds, however, there are interactions between initial biases and later experience.

5.4 Effects of development on language and birdsong typology

What can vocal development tell us about the evolution of sounds? Cultural evolution of signals is only possible by virtue of learning. Even though humans are ‘open ended’ learners, meaning learning is still possible in adulthood, the major part of language learning occurs early in development. It is thought that humans and zebra finches have a sensitive phase during which they have to be exposed to language for proper language acquisition to be possible. Moreover, biological aspects of development may influence the learning process during this stage and affect the end product: the actual speech or song produced in adulthood, which will be the input for the next generation. This implies that it is highly relevant to study vocal development from an evolutionary perspective too. Studying the effects of experience dependent and independent processes on the adult end product, could shed light on the interaction between cultural evolution and biological processes.

The experience independent biases in songbirds and human infants may affect the final vocal productions of song and speech by constraining learning to a subset of the signal. The bias for homorganic patterns in infants seems to affect at least later child language productions, since homorganic syllables predominate in babbling and first words (Davis & Macneilage, 1995; Fikkert & Levelt, 2008; Levelt, 1994; Levelt, 1995). As described above, it is unclear yet whether the origin of the bias lies in perceptual or motor mechanisms or both, it is most likely independent of auditory experience. Previously, MacNeilage and Davis have suggested that CV-co-occurrence (similar to homorganicity) is also widespread across adult languages. These authors mention that historical linguistic research also indicates that these homorganic (or CV-co-occurrence) patterns may be part of a 'protolanguage' (MacNeilage & Davis, 2000). This may have an evolutionary background in sucking and chewing (MacNeilage, 1998a; MacNeilage, 1998b), resulting in repetitive jaw movements. These movements, together with phonation, form the first babbling sounds. According to MacNeilage and Davis this could have led to the common occurrence of such patterns in present language. The bias for homorganicity may thus have effects not only on child language but also on adult language. The present data demonstrate that there is an (additional) perceptual bias for homorganicity. The proposed account, an initial suprasegmental representation, might explain the findings on adult language data as well; even though adults have a segmental representation, if language acquisition starts out with a suprasegmental representation, this may increase the chance that homorganic syllables arise in language. The claim by MacNeilage and Davis that homorganicity is frequent across languages, however, has been challenged by Albano et al. (2011). Moreover, heterorganic patterns are also observed cross-linguistically, which complicates the matter. The analyses showing higher frequency cross-linguistically for homorganic patterns (or CV-co-occurrence) are based on observed to expected ratios within language, a calculation that has not been used for more general typological analyses. This aspect needs to be further investigated to better understand the relation between developmental patterns and cross-linguistic patterns.

In zebra finches, a direct effect of the initial bias on adult vocalizations is shown. Birds tutored with more common songs (for which they have an initial bias) produced songs that more accurately resembled the tutor song. This shows there is also a bias for producing these common types of elements, resembling what is found in child language. One study on zebra finch song learning in naive birds, does suggest that biases can lead cultural evolution in a certain direction (Feher et al, 2009). In this study, naive birds were used as tutors for other naive birds, which were thus tutored with isolate songs. This iterative learning was repeated for a few 'cultural generations'. After approximately 4 of these generations the birds' songs were similar to wild type song and no longer resembled isolate song. This suggests that predispositions guide the learning towards wild type song. The initial bias found in chapter 3 might be one of those biases affecting song production patterns.

Possibly this can lead to increasing frequency of occurrence and/or stabilization of such preferred elements within and across populations. If this is indeed the case, the effect should become evident in Chapter 4, where the distribution of elements within and across populations is analyzed. Chapter 4 shows how elements are distributed across 12 laboratory populations and one wild population. It is difficult to directly compare the types of elements classified as common or uncommon in chapter 3 (the juvenile zebra finch experiment) with element clusters in chapter 4, since the classification in chapter 3 was based on visual interpretation of the spectrograms and the one in chapter 4 by means of acoustic analyses. Thus it is not known whether the elements classified as common (chapter 3) would match certain clusters of the acoustic analysis (chapter 4). In chapter 3, high notes (especially inspiratory ones), trills, high sweeps and noisy elements are classified as uncommon. Subjectively, these are most comparable with the clusters 8, 3, 11, respectively, for the first three, while 1, 6 and 10 possibly resemble noisy elements (see fig 4.3a in Chapter 4). The elements classified as more common in Chapter 3 are stacks, slides, short slides and tones, which are probably covered by the clusters 2, 4, 5, 7 and 9 (see fig 4.3a in Chapter 4). Objective acoustic analyses should reveal how well this subjective classification fits. Regardless of this classification, it is clear that none of the elements really dominates. On the other hand, none of the elements is absent from any of the populations either, indicating that in this analysis all elements are 'universal'. There is variation between populations in how the elements are distributed across populations, in the sense that some elements are more abundantly present in one population than in the other. This variation in distribution, however, does not show a clear relation to the classification of elements into universal or non-universal used for the juvenile zebra finch stimuli. Therefore, it is at this point unclear to what extent the early bias affects sound patterns across and within populations. It should be taken into account that acoustic and clustering analyses were performed on elements of all populations together. Slightly different patterns might occur when analyses would be performed on each population separately. For instance it is possible that in the range of one cluster in the present analysis, two clusters are formed in a one-population analysis (fig. 5.1).

This makes the comparison between this analysis and a typological linguistic analysis more difficult as well. Phonological typology in human language is usually based on categories that can be verified by speakers of that language. Unfortunately, we cannot ask the birds directly to verify which categories they use, and as far as we know the elements in birdsong are not contrastive. Discrimination and categorization tests in songbirds have been performed previously (Prather et al, 2009; Sturdy et al, 1999a; Vicario et al, 2001). A study on zebra finches revealed that they have a categorization that is somewhat comparable to the classification used by Sturdy et al. (Sturdy et al, 1999b), which is again different from the clusters found in Chapter 4. (Although no evidence for categorical perception was shown). This direction of research should be further explored to find out how the clusters found in chapter 4 are perceived by zebra finches.

Apart from how birds categorize song, it is hard to predict what an acoustic clustering analysis similar to the one in Chapter 3 would result in when performed on human speech. It is likely that some clusters will be found when analyzing sound pronunciations using phonological or phonetic information in human speech (Miller, 1989). For instance, clustering has been shown for vowel distribution, at least within a particular language. The distribution into clusters with high frequency in the centre of the cluster and low frequency around the borders of the cluster (bimodal distribution) has been thought to be a possible cue for infants to learn phonetic categories. Usually however, these comparisons are based on the features known to be the cue for a segmental contrast (for instance formant frequency (Miller, 1989)). Furthermore, phoneme clusters seem to overlap too, even within a language, thereby blurring cluster boundaries. Nevertheless, some language-specific features are so different from others (for instance click sounds (Herbert 1990)) that it seems likely they would form a cluster that does not occur in other languages. If this is correct, it follows that there is more diversity among languages than among songs in different zebra finch populations.

In summary, early perceptual biases affect production patterns later in development in songbirds as well as in human children. Whether this effect of the bias on song and speech production results in patterns cross-linguistically or across songbird populations is not clear yet.

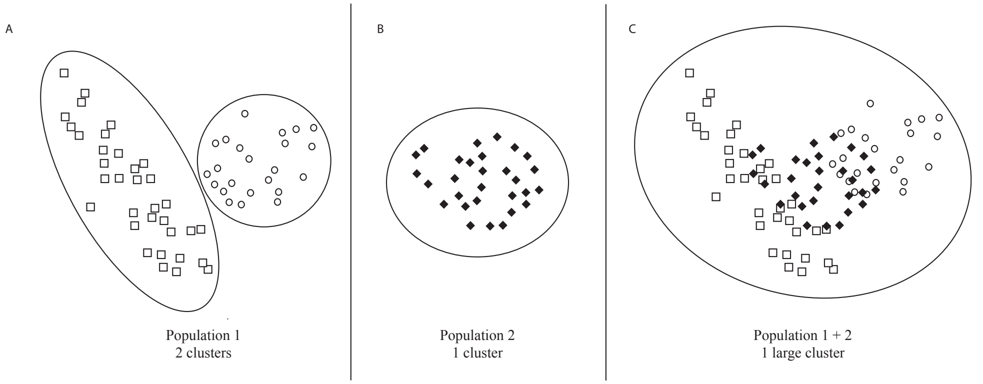

Figure 5.1. A schematic representation of how grouping populations in the analysis might blur some cluster boundaries. If a separate analysis on population 1 reveals 2 clusters (A) and a separate analysis of population 2 reveals one (intermediate) cluster, an analysis where population 1 and 2 are analyzed together might blur the cluster boundaries and group all three clusters into one large cluster (C).

5.5 Conclusion

This thesis has demonstrated that phonological development in both human infants and in zebra finches is affected by factors dependent on, as well as independent of adult input. Early biases have been shown in infants to be independent of frequency of occurrence, and in zebra finches biases were present before song exposure. Such biases can guide the learning process early in development. Whether these biases have an effect on the sound patterns within and between populations and can explain mechanisms like sound change and/or evolution is at this point unclear. No direct relation is found between the early biases in zebra finches and population wide acoustic clustering analyses, whereas in chapter 3 the birds' songs probably were affected by the initial bias. Further research is needed to elucidate this issue. What the acoustic clustering analysis does show is that, similar to language, population differences are found in the way elements are distributed, some elements occurring more often in one population than in the other.

The influence of experience dependent factors is very clear in the study on juvenile zebra finches where the type of elements birds are tutored with, determines which elements they prefer to listen to after tutoring. The exposure also affects their song copying (as expected) but the copying is slightly more accurate in birds tutored with common elements (for which the birds have an early bias) indicating an interaction between the early bias and subsequent exposure. In infants the effect of experience is less clear, since no effect of frequency was found. The preference changes over the course of development as well, but if this change is driven by experience or because of maturation is at this point unclear. Obviously infants are influenced by experience in one way or another, so an interaction between experience dependent and independent processes is likely to explain these results. The perceptual data fit the child language production data very well, which also shows an early preference for homorganic syllables, while heterorganic syllables occur only later. Exploring the interaction between experience dependent and experience independent processes in vocal learning in more detail is a challenge for future research.