Perception & Psychophysics
2008, 70 (8), 1515-1525
doi: 10.3758/PP.70.8.1515

The quest for generalizations over consonants: Asymmetries between consonants and vowels are not the by-product of acoustic differences

Juan M. Toro
Scuola Internazionale Superiore di Studi Avanzati, Trieste, Italy
and Universitat de Barcelona, Barcelona, Spain

Mohinish Shukla
University of Rochester, Rochester, New York

Marina Nespor
Università di Milano Bicocca, Milan, Italy
and Centro Linceo Beniamino Segre, Rome, Italy

and

Ansgar D. Endress
Harvard University, Cambridge, Massachusetts

Consonants and vowels may play different roles during language processing, consonants being preferentially involved in lexical processing, and vowels tending to mark syntactic constituency through prosodic cues. In support of this view, artificial language learning studies have demonstrated that consonants (C) support statistical computations, whereas vowels (V) allow certain structural generalizations. Nevertheless, these asymmetries could be mere by-products of lower level acoustic differences between Cs and Vs, in particular the energy they carry, and thus their relative salience. Here we address this issue and show that vowels remain the preferred targets for generalizations, even when consonants are made highly salient or vowels barely audible. Participants listened to speech streams of nonsense CVCVCV words, in which consonants followed a simple ABA structure. Participants failed to generalize this structure over sonorant consonants (Experiment 1), even when vowel duration was reduced to one third of that of consonants (Experiment 2). When vowels were eliminated from the stream, participants showed only a marginal evidence of generalizations (Experiment 4). In contrast, participants readily generalized the structure over barely audible vowels (Experiment 3). These results show that different roles of consonants and vowels cannot be readily reduced to acoustical and perceptual differences between these phonetic categories.

Children have to acquire many different aspects of their native languages. These aspects can be roughly classified in two categories, one pertaining more to learning words and their meanings, and one pertaining more to acquiring grammatical aspects of language. Of course, word learning depends at least in part on grammatical information (e.g., Gillette, Gleitman, Gleitman, & Lederer, 1999), and grammatical regularities may partly depend on lexical information (e.g., Tomasello, 2000). Still, different cues in the speech signal may be used in different ways for acquiring specific parts of language. In particular, Nespor, Peña, and Mehler (2003) suggested on theoretical grounds that consonants might be more relevant for lexical processing, whereas vowels may have a more grammatical function.

The question whether vowels and consonants have different functional roles is related to a more general issue. Traditionally, language acquisition has been thought to be possible only due to strong (probably innate) biases that shape how linguistic stimuli are processed (Chomsky, 1980). More recently, however, different authors have proposed that much of language acquisition can be accounted for by more general mechanisms that operate in a variety of domains and exploit distributional regularities in their input (Elman et al., 1996; McClelland, Rumelhart, & the PDP Research Group, 1986). Although functional asymmetries between vowels and consonants are compatible with the former approach (because vowels and consonants would be intrinsically linguistic categories), one would not expect such differences if grammar were learned exclusively through general learning mechanisms, because, all else being equal, either stimulus should be equally good for allowing such learning.

A. D. Endress, ansgar.endress@m4x.org

 Copyright 2008 Psychonomic Society, Inc.

Experiments using artificial languages have suggested that consonants are indeed preferentially used for identifying words (Bonatti, Peña, Nespor, & Mehler, 2005; but see Newport & Aslin, 2004),[1] whereas vowels are used for extracting simple grammar-like rules (Toro, Nespor, Mehler, & Bonatti, 2008). Although the authors of this latter study interpreted their results as evidence that vowels are more important for learning morphosyntactic grammar-like regularities than consonants are, there is a simple alternative explanation. Indeed, the "rules" implemented in that study were based on repetitions of vowels embedded in artificial nonsense words (for example, ***ta*pen*a***, where the vowel of the first syllable is the same as the vowels of the last syllable; hereafter, we will call such structures *repetition-based*). However, such simple repetition-based structures are learned preferentially when they are made highly salient; in fact, when such structures are presented in less salient ways, participants failed to learn them (Endress, Scholl, & Mehler, 2005). It is thus possible that participants preferred to use vowels for learning these rules, not because vowels are particularly suitable for learning grammatical regularities, but rather because they carry much more energy and are much more salient than consonants (e.g., Ladefoged, 2001, 2006). In other words, repetition-based structures may be generalized preferentially on the most salient items available, and thus, in a sequence of consonants and vowels, on the vowels; the fact that vowels are a linguistic category may be a mere coincidence.

Here, we investigate the question whether generalizations are preferentially observed over vowels because of their specific linguistic functions, or whether the asymmetries observed in earlier experiments were just a side effect of the vowels' salience. We thus attempt to obtain generalization over consonants by progressively increasing their salience, and by simultaneously decreasing the salience of the vowels. Whereas we find reliable rule learning with vowels even under the most extreme conditions, we do not—to anticipate our results—find significant learning of consonant-based rules in any of the experiments. Still, when pooling together all experiments, or removing vowels altogether, a trend toward generalization over consonants emerges. Hence, rules can also be learned over consonants; however, all things being equal, it is much easier to learn rules over vowels.

The Possible Functions of Vowels and Consonants

What is the nature of the processes involved in language acquisition and use? At their extremes, two answers have been given to this question. On the one hand, a general mechanism that picks even subtle regularities in the speech signal may eventually give rise to a fully developed linguistic system with all its complexities. On the other hand, different specialized mechanisms may be needed to extract information from different sources in the signal. This may be true for broad distinctions, such as those between learning words and grammatical regularities (Pinker, 1999), but even within the domain of grammar, many different specialized mechanisms may conspire to give rise to a functional language system.

Studying the functional differences between vowels and consonants is a particularly attractive domain in which the general question can be tested, because it can easily be experimentally manipulated. If language fundamentally relies on a single, general, unitary system, both consonants and vowels should fit equally well as input for linguistic computations; there is, therefore, no need to presuppose representational differences between them. Alternatively, these phonological representations might convey distinct types of information, and may trigger different kinds of specialized computational processes.

From a linguistic perspective, one might indeed expect vowels and consonants to carry different types of information during language processing. Whereas consonants are more readily engaged in lexical processing, vowels play a primary role in marking syntactic constituency (Nespor et al., 2003). Indeed, it is well known that, across different languages, lexical information relies predominantly on consonants. This is particularly apparent in Hebrew and several other Semitic languages, where lexical roots are made up only by consonants and where, in writing, letters code mostly for consonants (whereas vowels may optionally be specified by diacritics). A related effect is found experimentally in adult participants. When they have to change a phoneme to transform a nonword into a real word, they preferentially replace a vowel, and not a consonant; for example, they are more likely to change *cebra* to *cobra* (a vowel change) rather than to *zebra* (a consonant change; Cutler, Sebastián-Gallés, Soler-Vilageliu, & van Ooijen, 2000; Sharp, Scott, Cutler, & Wise, 2005; van Ooijen, 1996), suggesting that consonants constrain lexical access more strongly.

Vowels, on the contrary, mainly carry prosodic information through pitch changes, or lengthening, among other suprasegmental features. Since this information can signal aspects of syntactic structure (Nespor & Vogel, 1986), vowels provide the listener with cues about how units, such as individual words, are to be organized in language. Thus, prosodic cues allow infants to learn important syntactic regularities (Christophe, Nespor, Guasti, & van Ooyen, 2003; see also Gleitman & Wanner, 1982; Morgan & Demuth, 1996), although direct evidence for such a role of the vowels is still lacking.[2]

Of course, these differences between vowels and consonants are relative rather than absolute. Many words differ only in their vowels (e.g., minimal pairs such as *pan*, *pen*, and *pin*, or *ball*, *bell*, and *bill*); if vowels were ignored for lexical processing, it would be impossible to keep these words apart. Likewise, consonants contribute to signaling syntactic constituency in some cases, such as the French *liaison*. In French, the final (usually unpronounced) consonant of a word gets resyllabified with the initial vowel of the next word (in case it starts with a vowel) if the two words belong to the same phrasal constituent. For example, in *les ours ont mangé* ("the bears have eaten"), a liaison occurs between *les* and *ours* because these words belong to the same constituent, but not between *ours* and *ont* because these words have a lower level of syntactic cohesion (technically, liaison occurs

only within phonological phrases; see Nespor & Vogel, 1986; Selkirk, 1974). Hence, we do not suggest that there is no lexical role for vowels, or no grammatical role for consonants, but rather that, all things being equal, vowels are more important for signaling constituent structure, whereas consonants are more important for signaling lexical items.

Dissociations Between Vowels and Consonants

The notion of a functional difference between consonants and vowels has been strengthened by a growing number of experimental results. Drawing from neuropsychological case studies, Caramazza, Chialant, Capasso, and Miceli (2000) reported selective impairment of consonants and vowels in two patients. Whereas one of the patients presented important difficulties only in processing vowels, the other patient presented difficulties only in processing consonants, suggesting that different neural substrates might be involved in the processing of the two phoneme classes. Importantly, such a selective impairment is not due to a failure in processing lower level acoustic features that differentiate consonants from vowels (e.g., those correlated with the fact that vowels are sonorant); it is unclear, therefore, how the dissociation can be explained without relying on a categorical distinction between the two types of linguistic representations (but see Monaghan & Shillcock, 2003, 2007, and the response by Knobel & Caramazza, 2007).

Evidence for an important role of consonants in lexical processing also comes from laboratory experiments with infants. Nazzi (2005), for example, demonstrated that consonants, not vowels, play a primary role in word learning for infants at 20 months of age; but see Mani and Plunkett's (2007) investigation of the perception of mispronounced known words. Importantly, such an effect is found even when vowels are compared with either plosive or nonplosive consonants (Nazzi & New, 2007), ruling out the possibility that these results are just produced by differences in sonority levels.

An advantage for consonants over vowels was shown, even with adults, in statistical learning experiments designed to model word segmentation. In such experiments, participants are typically exposed to a continuous speech stream comprising a series of nonsense words. Thus, the only cue to word boundaries is that syllables within words are more likely to follow each other than are syllables across words; that is, the "transition probabilities" (TPs) are higher within words than between words (Aslin, Saffran, & Newport, 1998; Saffran, Aslin, & Newport, 1996).[3] Bonatti et al. (2005) showed that adult learners preferentially perform such statistical computations over consonants than over vowels. After familiarization with a continuous speech stream, participants could effectively segment the words when the only reliable cue was the TPs between the consonants forming them. In contrast, participants could not segment the stream when they had to rely on statistical information between vowels, unless the information was highly redundant (i.e., when only two frames were used for creating the words, so the stream contained immediate repetitions; see also Newport & Aslin, 2004). Moreover, when faced with misaligned consonantal and vocalic information, participants rely more heavily on the former than on the latter (Mehler, Peña, Nespor, & Bonatti, 2006). Thus, just as consonants seem to be important for lexical processing, they are also the preferred target representations over which words are extracted from continuous speech.

Other authors have argued that these results were mere by-products of acoustical or distributional differences between vowels and consonants in the participants' native language, and not the result of different processing by the linguistic system. For example, Keidel, Jenison, Kluender, and Seidenberg (2007) claimed that differences in the distribution of consonants and vowels in French (the native language of the participants in Bonatti et al., 2005, and Mehler et al., 2006) may explain the above-mentioned results. They showed through corpus analyses that, at least in French, consonant sequences are better predictors of lexical identity than are vowels.[4] Adults may have learned through years of experience that consonants are more significant for lexical recognition than are vowels; these differences could explain why consonants were the preferred target of statistical computations during the artificial language experiments.

Still, there is no evidence that the direction of causation is the one proposed by Keidel et al. (2007) and not the opposite one. After all, one has to explain where the distributional differences come from; if they are due to the processing biases proposed in the aforementioned experiments, one has a principled explanation for both the experimental and the distributional results. By Keidel et al.'s account, Bonatti et al.'s (2005) results may be explained, but the cross-linguistic distributional differences would be a mere accident, as is touched on in the General Discussion (see Bonatti, Peña, Nespor, & Mehler, 2007, for more details).

Recent experiments with Italian speakers have shown further evidence for a complementary functional asymmetry between vowels and consonants (Toro et al., 2008). In their experiments, the authors presented participants with an artificial language containing words formed by statistically coherent "consonantal frames" and vowels following a simple underlying structure. After 10 min of presentation of this language, participants readily extracted the words using the consonantal information (as in Bonatti et al., 2005), and generalized the simple rule implemented over the vowels. In contrast, when a different set of participants was presented with a "mirror" artificial language, in which the roles of vowels and consonants were inverted (that is, the statistically coherent frames were implemented over vowels and the underlying structure over consonants), they could extract neither the (statistically defined) words nor the structure. These results confirm the role of consonants as a target of statistical computations used for lexical processing, and suggest a preferential use of vowels for some structural generalizations.

However, one might ask what the source of such differences might be. Consonants and vowels differ in a number

of acoustic parameters (e.g., Clements, in press; Ladefoged, 2001, 2006). A very salient acoustic difference between consonants and vowels is that vowels carry more energy than consonants; vowels are thus more salient even to babies (Mehler, Dupoux, Nazzi, & Dehaene-Lambertz, 1996). Therefore, it may be argued that, whatever element in the speech signal carries more energy or is more salient within a given sequence will be privileged for the extraction of structural generalizations. This is particularly true for the sequences used by Toro et al. (2008). Indeed, their rules were based on repetitions of vowels (or consonants); more specifically, the first vowel in a word had to be the same as the last one (as in *t**a**pen**a***).

Repetitions, however, may not be representative of rule learning in general (although there certainly are repetition-based regularities, for instance in Semitic languages; McCarthy, 1979); rather, repetitions seem to be processed by a specialized repetition-detecting "primitive" (Endress, Dehaene-Lambertz, & Mehler, 2007; Gómez, Gerken, & Schvaneveldt, 2000; Tunney & Altmann, 2001). Moreover, repetition-based structures are generalized only under particularly suitable conditions (Endress et al., 2005). It is thus possible that the repetition-based structures in Toro et al.'s (2008) experiments were also extracted just on the most salient items available, and these happened to be the vowels. If so, the asymmetries observed by Toro et al. may simply be due to the fact that vowels were more salient, with no deep implications for the nature of the representations used to learn more lexical or more grammatical aspects of language

The Present Experiments

The aforementioned considerations suggest that the observed difference in the roles of consonants and vowels may be reduced to a difference in terms of the energy carried by different elements in the speech signal (and thus their saliency). In the present article, we address this possibility by investigating how easily participants would generalize simple structures over consonants that carry more energy than vowels. If the energy hypothesis is correct, we should observe structural generalizations over consonants just as they have been observed over vowels. This issue was investigated in four experiments. As a first step, we created an artificial speech stream in which consonants conformed to a simple rule. Importantly, only sonorant consonants (like /n/; that is, consonants that are as audible as vowels) were used in this experiment. Such consonants should carry a higher amount of energy than the ones used in previous experiments (because the consonants in the previous experiments were to a large extent stops, such as /t/). In Experiment 1, we investigated whether participants could generalize a simple structure over sonorant consonants. In Experiment 2, we made consonants even more salient by shortening vocalic segments to one third of the duration of consonants, and by concurrently lengthening the consonants. In Experiment 3, the simple structure was implemented over the reduced vowels, to assess participants' generalization abilities over very reduced vowels. Finally, in Experiment 4, vowels were completely eliminated from the speech stream, to test whether, under this radical modification, participants could effectively generalize the structure to new tokens on the basis of consonantal information.

EXPERIMENT 1
Sonorant Consonants

In this experiment, we familiarized participants with a speech stream in which words were defined over the vocalic tier, while consonants implemented a simple structure. Importantly, all consonants employed during familiarization were sonorants. After familiarization, participants had to complete two kinds of test trials. In the "recognition" trials, we simply assessed whether participants recognized the items they had heard. In the "generalization" trials, we asked whether participants would generalize the regularity implemented by the consonants to new tokens. Importantly, in the present article we do not directly address the question of whether TPs can be computed over vowels, since other researchers have already explored this issue (Bonatti et al., 2005; Newport & Aslin, 2004). In our experiments, we included small pauses between words; this made the recognition of individual words possible without the need to rely on statistical computations. Here, we focused on how feasible generalizations over consonants are. Above-chance results in the generalization test would show that using the more salient sonorant consonants is enough for allowing participants to draw structural generalizations over them.

Method

Participants. Participants were 15 native speakers of Italian, all undergraduate students. None reported hearing deficits; they were paid for their participation in the study.

Stimuli. Stimuli were 12 trisyllabic nonsense words. Each syllable had a consonant–vowel (CV) structure. In the following, we will call the phoneme sequences carrying statistical information the *word frames*, and the phoneme sequences implementing the structural regularity the *slots*. In Experiment 1, the word frames contained vowels, and the slots carried consonants (because we implemented the structural regularity over consonants). The words were thus created over two vocalic frames (a_E_u; i_o_e), and three sonorant consonants (n, l, R) were used for filling the consonant slots. Importantly, these consonants followed an ABA structure (e.g., ***R**a**N**e**R**u*; see Table 1 for the full list). Words were concatenated to create a

Table 1
Vocalic Frames and Consonantal Structures Used to Compose Words in Experiments 1 and 2

Consonantal Filler	Vocalic Frame	Word
R_n_R	a_E_u	RanERu
R_l_R		RalERu
n_R_n		naREnu
n_l_n		nalEnu
l_R_l		laRElu
l_n_l		lanElu
R_n_R	i_o_e	RinoRe
R_l_R		RiloRe
n_R_n		niRone
n_l_n		nilone
l_R_l		liRole
l_n_l		linole

continuous speech stream. In order to avoid immediate repetitions of vocalic frames, one to three syllables were inserted between the words. These syllables were identical to the ones composing the words, but they were combined randomly, so they did not exhibit any structure that could be extracted. In addition, 25-msec pauses were inserted before and after each word, in order to facilitate the identification of the vocalic frames.

The stream was then synthesized with MBROLA (Dutoit, Pagel, Pierret, Bataille, & van der Vrecken, 1996), using an Italian female database (it4). F0 for all phonemes was set to 240 Hz, and their durations were set to 120 msec. Test items were synthesized with the same parameters. For the recognition test, words and part words were created. Part words were made of vowels spanning two vocalic frames (as in the "part frame" u_i_o, which is constructed from the last vowel from frame a_E_u and the first two vowels of frame i_o_e), and were filled with the same consonants, following the ABA structure as in the words. Vocalic information was thus the only cue for differentiating these test items.

For the generalization test, rule words and nonrule words were created. Rule words were identical to the words, but two new consonants (j, m) were used for filling the consonantal gaps. Nonrule words were identical to rule words, but the two new consonants followed either an AAB or an ABB, not an ABA, structure. These were equally represented in the test pairs. Consonantal information was thus the only reliable cue for a correct discrimination among these test items.

Procedure. Participants were tested individually in a silent room, wearing headphones. An Apple G4 running PsyScope X (psy.ck.sissa.it) controlled the experiment. Participants were told that they would hear an artificial language, and that their task was simply to listen to it. After 10 min of familiarization with the speech stream, participants were informed that they would hear pairs of words, and that they would have to decide which word in each pair was more likely to be from the previously heard language. Then, they completed an auditory two-alternative forced choice test, with two kinds of test pairs: words versus part words, and rule words versus nonrule words. The first comparison (*recognition test*) assessed the participants' ability to remember elements from the stream when relying only on vocalic information. The second comparison (*generalization test*) tested their ability to extract the consonantal structure. The 16 test trials (8 for each comparison) were interleaved in semirandom order, with the restriction that no more than two trials of the same type could occur successively. In each trial, test items were separated by a 500-msec pause.

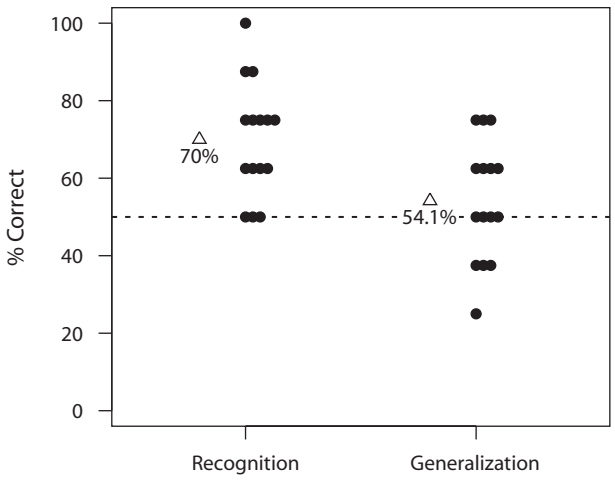

Figure 1. Individual scores and group means during the recognition and generalization tests in Experiment 1. Dots represent the means of individual participants; triangles, sample averages; and the dotted line, the chance level of 50%. When familiarized with a speech stream in which simple rules are carried by sonorant consonants, participants fail to generalize the rules, but recognize the items they have heard.

Results and Discussion

As shown in Figure 1, participants correctly recognized the words in the recognition test well above chance (M = 70%, SD = 14.7%) [$t(14) = 5.23$, $p < .001$]. In the generalization test, in contrast, they had no preference for rule words over nonrule words (M = 54.1%, SD = 15.4%) [$t(14) = 1.04$, $p < .313$, n.s.].[5] The use of sonorant consonants (that might be expected to be more salient) in the present experiment thus did not improve the generalization of the structural information to new tokens. This suggests that observed differences between consonants and vowels in the present article may not just be a by-product of differences in their relative salience, but most likely a consequence of differential functional roles played by these elements.

Still, one could argue that, compared with sonorant consonants, vowels carry more total energy, so the energy hypothesis has not been ruled out. Indeed, we analyzed 5-sec samples of the stimuli from Experiment 1, using PRAAT. Onsets and offsets of individual vowels and consonants were marked manually, and the energy computed using the "get energy" function in PRAAT—the energy being defined as the integral of the square of the amplitude between two time points.[6] We found that, on average, vowels still carried 1.78 times the energy of consonants. To use a stronger reversal of the energy relation between consonants and vowels, in the next experiment we reduced vowel duration to one third of the duration of each consonant. At least with this modification, consonants should be perceptually more salient than vowels.

EXPERIMENT 2
Reduced Vowels

In this experiment, we extended the results obtained in Experiment 1 to assess whether participants can track a structural regularity implemented over consonants. In addition to using relatively salient consonants, as in Experiment 1, we added 60 msec to the duration of each consonant, and halved the vowel duration. By any account, consonants should thus be much more salient than vowels, which might facilitate extracting the structural regularity.

Method

Participants. Participants were 15 native speakers of Italian, all undergraduate students. None reported hearing deficits; they were paid for their participation in the study.

Stimuli. The stimuli were identical to those of the previous experiment, with one exception. The duration of all vowels was reduced from 120 to 60 msec, and the duration of all consonants was increased from 120 to 180 msec. Duration of vowels in all test items was also only one third of that of the consonants. With this modification, the relation between consonants and vowels, in terms of total energy, was clearly inverted. Whereas vowels carried 1.78 times the energy of consonants in Experiment 1, consonants carried 1.34 times the energy of vowels in Experiment 2.

Procedure. The procedure was identical to that of the previous experiment.

Results and Discussion

As shown in Figure 2, participants correctly recognized the words over the part words ($M = 66.6\%$, $SD = 17.4\%$) [$t(14) = 3.69$, $p < .005$]. However, they did not show any preference for rule words over nonrule words in the generalization test ($M = 55\%$, $SD = 11.3\%$) [$t(14) = 1.7$, $p < .111$, n.s.]. These results closely replicate those from Experiment 1, since an ANOVA using experiment (Experiment 1 vs. 2) as a between-subjects factor and type of test (recognition vs. generalization) as a within-subjects factor showed neither a significant difference between experiments [$F(1,28) = 0.145$, $p = .706$, n.s.] nor a significant interaction between experiments and type of test [$F(1,28) = 0.229$, $p = .636$, n.s.].

Even when consonants were three times as long as vowels, no generalization over consonants emerged. The preferential use of vowels to draw structural generalizations over consonants thus does not seem to depend on their relative salience within the speech stream.

In order to clarify the asymmetric roles of vowels and consonants, we decided to evaluate another prediction: If the advantage for vowels is due to their being more salient than the consonants, one would expect a decline in the generalization performance when this salience difference is reduced. To assess this possibility, we created a new stream inverting the role played by consonants and vowels. In the new stream, statistical frames were implemented over consonants, and the vowels filled the intervening slots, following the ABA structure. In this new stream, we kept the length of the vowels to one third of that of the consonants, so the former would be much less salient than the latter.

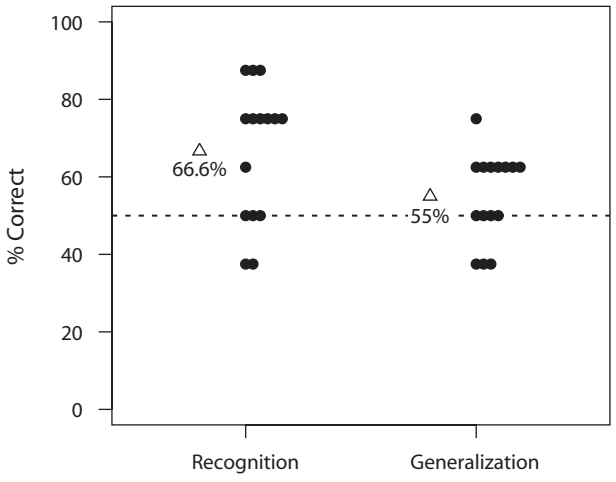

Figure 2. Individual scores and group means during the recognition and generalization tests in Experiment 2. Dots represent the means of individual participants; triangles, sample averages; and the dotted line, the chance level of 50%. When familiarized with a speech stream in which a simple rule is carried by sonorant consonants, participants fail to generalize the rule, but recognize the items they have heard even when the vowel duration is reduced to one third of the consonant duration.

Table 2
Consonantal Frames and Vocalic Structures Used to Compose Words in Experiment 3

Vocalic Filler	Consonantal Frame	Word
a_e_a	n_l_s	nalesa
a_o_a		nalosa
e_a_e		nelase
e_o_e		nelose
o_a_o		nolaso
o_e_o		noleso
a_e_a	j_R_m	jaRema
a_o_a		jaRoma
e_a_e		jeRame
e_o_e		jeRome
o_a_o		joRamo
o_e_o		joRemo

EXPERIMENT 3
Generalization Over Reduced Vowels

In Experiment 2, participants failed to generalize a structural regularity carried by consonants although salient (sonorant) consonants were used that were three times as long as the vowels. By any account, this manipulation should encourage the processing of consonants (and discourage the processing of vowels). Still, participants failed to learn the structural regularity, and did not perform any better than in Experiment 1. In Experiment 3, we kept the reduced vowels (and augmented consonants) from Experiment 2, and asked whether generalizations over vowels are possible, even under conditions where the vowels are impoverished.

Method

Participants. Participants were 15 native speakers of Italian, all undergraduate students. None reported hearing deficits; they were paid for their participation in the study.

Stimuli. Stimuli were created as in the previous experiment, except that the frames were implemented over consonants (n_l_s, j_R_m), and the intervening slots were filled with vowels (a, e, o) following the ABA structure (see Table 2); that is, in this experiment the statistical information was carried by the consonants, whereas the structural regularity was implemented by the vowels. Importantly, the duration of all vowels was kept at 60 msec, but the duration of consonants was 180 msec. Small (25-msec) pauses were inserted between words. In this way, the present stream closely mirrors that used in Experiment 2, with the difference that statistically coherent frames were implemented over the consonants, and the structures to be generalized were implemented over the vowels. During the generalization test, we used two new vowels (I, u) that did not appear during familiarization.

Procedure. The procedure was identical to that of the previous experiments.

Results and Discussion

As shown in Figure 3, participants readily segmented the words from the stream using consonantal information; they thus performed well on the recognition test ($M = 69.1\%$, $SD = 13.2\%$) [$t(14) = 5.60$, $p < .005$]. In a marked contrast to the previous experiments, participants performed reliably above chance even in the generalization test ($M = 61.6\%$, $SD = 12.9\%$) [$t(14) = 3.50$, $p < .005$], even though the vowels were reduced to one third of the consonant duration, and were barely audible. Par-

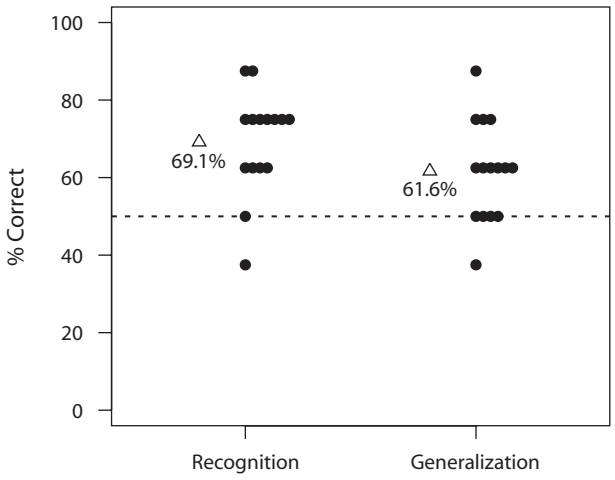

Figure 3. Individual scores and group means during the recognition and generalization tests in Experiment 3. Dots represent the means of individual participants; triangles, sample averages; and the dotted line, the chance level of 50%. When familiarized with a speech stream in which a simple rule is carried by vowels whose duration is reduced to one third of that of the consonants, participants readily generalize the rule. They also recognize the items they have heard.

ticipants thus reliably generalize simple structures when these are implemented over vowels.

It is also interesting to compare the results of Experiments 2 and 3. Since these experiments differ in two variables—namely, the carriers of the generalizations (consonants in Experiment 2; vowels in Experiment 3) and the saliency of these carriers (maximized in Experiment 2; minimized in Experiment 3)—it is difficult to attribute differences between these experiments to any of these variables. Still, even though the results of Experiments 2 and 3 did not differ significantly [$F(1,28) = 2.066$, $p = .162$, n.s.], the effect size in the generalization condition of Experiment 3 was more than twice as large as the one in Experiment 2 (Cohen's $d = 1.278$ for Experiment 3; Cohen's $d = 0.621$ for Experiment 2). Hence, even when every effort was made to discourage generalizations on vowels and encourage generalizations on consonants, we observed robust generalizations on vowels but not on consonants. Still, the lack of a significant difference between these two experiments also supports the conclusion we will draw below—namely, that the difference between vowels and consonants in their ability to support generalizations is not an all-or-nothing distinction. Rather, all things being equal, vowels are more suitable for such generalizations than are consonants.

In the final experiment, we tested an even more radical implementation of the energy difference between consonants and vowels. In this experiment, we completely eliminated vowels from the familiarization stream, leaving only consonantal segments. If difficulties for making generalizations over consonants resulted from a difference in the perceptual salience among phonological representations, participants should be able to easily generalize the simple structure when only consonantal segments are presented.

EXPERIMENT 4
Generalization Over Consonants With No Vocalic Segments

Experiments 1 through 3 showed that, whereas generalizations over vowels are readily made even when the vowels are greatly diminished perceptually, generalizations over consonants cannot be demonstrated, even when consonants are made highly salient. Here, we made a final attempt to observe generalizations over consonants by completely eliminating all vocalic segments from the speech streams.

Method

Participants. Participants were 20 native speakers of Italian; all were undergraduate students. None reported hearing deficits; they were paid for their participation in the study.

Stimuli. The stimuli were changed relative to the previous experiments because isolated consonants cannot be synthesized with MBROLA while avoiding coarticulation. A female native speaker of Italian thus produced the five consonantal segments (/r/, /n/, /l/, /m/, /s/) used in this experiment. We then cut out the steady portion of each consonant for a duration of 180 msec. The stream was constructed with the same structure as that in Experiments 1 and 2, including intervening elements between words. The consonants *R*, *n*, and *l* were used for the familiarization, and the consonants *m* and s were used for the generalization test. (The phoneme /j/, used in previous experiments, was replaced by /s/ because of difficulties in pronunciation; no recognition test was administered, because the vowels were removed from the stream.) All segments were recorded in isolation, using a Sony ECM microphone connected to a laptop computer via an M-Audio preamp USB audio interface; then they were concatenated to form the stream. All vowels from the sequence used in Experiments 1 and 2 were replaced by 60-msec silences. In order to make the rhythmical units corresponding to words recognizable by the participants, silences of 200 msec (rather than 25 msec) were inserted before and after each word. Since there was no vocalic information in this experiment, participants were not presented with the recognition test, which was always performed over the items (consonants or vowels) not used to learn the rules. Hence, participants completed only the eight trials of the generalization test.

Procedure. The procedure was identical to that of the previous experiments.

Results and Discussion

As shown in Figure 4, results from the generalization test were marginally different from chance ($M = 58.7\%$, $SD = 18.1\%$) [$t(19) = 2.15$, $p = .044$], even though it was necessary to run 5 more participants than in the other experiments;[7] that is, when only consonantal information was presented to the participants, they were able to generalize the structure. This result is not at odds with those of Experiments 1 and 2. Indeed, when pooling results from these experiments, a marginal preference for legal items emerges in the generalization test [$t(29) = 1.88$, $p = .070$]. Hence, it is indeed possible to draw some structural generalizations over consonants (as one might suspect, by considering morphology in language). Note, however, that these results are weak, and that one has to increase the sample size drastically to observe such a result (by 33% in Experiment 4 and by 100% in the pooled experiments). This contrasts markedly with the reliable generalizations over vowels observed in Experiment 3; so although it is to some extent possible to draw generalizations over con-

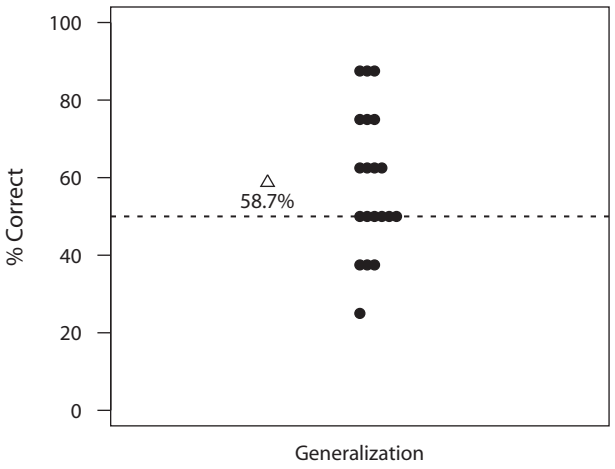

Figure 4. Individual scores and group mean during the generalization test in Experiment 4. Dots represent the means of individual participants; the triangle, the sample average; and the dotted line, the chance level of 50%. When familiarized with a sequence of sonorant consonants that carry a simple rule, participants tend to generalize the rule when vocalic information is completely removed from the sequence.

sonants, vowels are much favored, even under extreme conditions.

GENERAL DISCUSSION

A long-standing debate in language acquisition research is just how language-specific the underlying computations are. On the one hand, language may rely on a set of language-specific and largely innate abilities; on the other hand, language acquisition may rely in large part on more general, often statistical, learning mechanisms. In the context of this debate, an increasing number of studies have documented striking differences between the role of consonants and vowels. For example, studies with artificial languages have shown that consonants are preferred for the computation of statistical dependencies (Bonatti et al., 2005), whereas vowels seem to be preferred for structural generalizations (Toro et al., 2008).

The origins of these asymmetries between vowels and consonants, however, are far from settled. For example, the participants' experience with the distributional properties of their native language may lead to an advantage for consonants in lexical processing (e.g., Keidel et al., 2007; but see Bonatti et al., 2007). Likewise, the advantage for vowels in the structural generalizations observed by Toro et al. (2008) may be due to systematic acoustic differences between vowels and consonants. Indeed, vowels carry more energy than do most consonants, and thus are more salient; if the structures used by Toro et al. are computed preferentially on the most salient items available (for which there is evidence; see Endress et al., 2005), the "functional" difference between vowels and consonants may simply be due to the higher salience of vowels, with no deeper linguistic implications.

In the present study, we tackled this issue by modifying the saliency of vowels in the speech stream. We attempted to make consonants as salient as possible, and thereby to encourage generalizations over them (and simultaneously to discourage generalizations over vowels).

The results showed that when the simple structure was implemented over sonorant consonants, participants could not generalize it to new tokens (Experiment 1). This remained true even when vocalic segments were reduced to one third of the duration of that of the consonants (Experiment 2); that is, just modifying the perceptual salience of the phonemes by means of phonetic quality and duration is not enough to achieve rapid generalizations over consonants. In contrast, when the simple structure was implemented over vowels, participants generalized the structure to novel items even when the vowel durations were reduced to a third of the consonant durations (Experiment 3). In a final, more radical manipulation, we completely eliminated the vocalic contents from the speech stream. Under these conditions, the results revealed marginal evidence for structural generalizations over the remaining consonants (Experiment 4).

In sum, our results demonstrate that it is possible, to some extent, to generalize simple structures over consonants—for example, by eliminating all vocalic contents from the stream (or, as in the pooled analyses of Experiments 1 and 2, by doubling the number of participants). However, no such manipulation is required to observe robust generalizations over vowels; even when the vowel duration was reduced to a third of the consonant duration, participants readily generalized the underlying structural relation. All things being equal, there seems to be a genuine advantage for vowels for extracting certain structural generalizations, even when all efforts are made to discourage such generalizations.

Distributional Differences Between Vowels and Consonants

As mentioned in the introduction, some authors attributed processing differences between vowels and consonants to differences in the respective distributions of these phonetic categories (e.g., Keidel et al., 2007). We believe that such an account is unlikely to explain our results. First, since there are no regularities based on repetitions of linguistic elements in Italian (our participants' native language) similar to those used in our experiments, it is not clear what kinds of distributional differences could account for our results. Second, from a distributional perspective, it is difficult to explain why the vowel advantage for generalizations is so remarkably resilient to manipulations of the respective salience of vowels and consonants, to the point that participants generalize even over barely audible vowels. After all, one would not expect distributional knowledge to make perceptual factors entirely irrelevant; so we believe that it is reasonable to conclude that the vowel advantage we observe is not due to distributional differences.

More generally, distributional accounts (such as the one proposed by Keidel et al., 2007) do not seem to offer a principled explanation for the processing asymmetries between consonants and vowels (see also Bonatti et al., 2007). In fact, such accounts implicitly assume most of Bonatti et al.'s (2005) proposal: They have to assume that

consonants and vowels can be reliably categorized, and are represented on different tiers (so that either transitional probabilities or mutual information calculations can be performed *separately* on a tier). Moreover, these accounts also imply that listeners can actually perform mutual information calculations separately on vocalic tiers and consonantal tiers over the entire linguistic corpus, and that they can efficiently compare the results of such calculations. However, no direct empirical evidence has been provided for these assumptions.

Keidel et al.'s (2007) reliance on the learners' ability to track fine-grained distributional regularities present in language may reflect a widespread belief that the presence of statistical regularities in a listener's input automatically implies that the listener is able to track them. In fact, this is inconsistent with numerous results from animal cognition that demonstrated preferential learning of certain associations over others (e.g., rats easily associate visceral sickness with tastes, but not with places; conversely, they easily associate physical pain with places, but not with tastes; Garcia, Hankins, & Rusiniak, 1974). So, even though humans can compute certain statistics over given sequences of elements, it does not follow that they can readily compute *any* statistic over any set of elements. Of course, it is important to ask what kinds of distributional information are possibly contained in the signal, and could be used by an ideal observer. However, given that animals generally are not ideal observers when it comes to statistical regularities, it is equally important to study empirically what kinds of distributional information can actually be processed.

In the case of the processing differences between vowels and consonants, we believe that one also needs to provide a principled explanation of why consonants are cross-linguistically more important for lexical processing than vowels are, and of why vowels may be preferentially used for structural generalizations. This could be explained naturally if there were intrinsic processing differences between vowels and consonants, but if these processing differences were just side effects of distributional regularities, this cross-linguistic generalization would be little more than an accident. This, and the fact that in Italian there are no regularities entailing repetitions of items that could have given a distributional bias to our participants, suggest that the vowel advantage for generalizations we observed is not due to distributional knowledge of some sort, but rather reflects some genuine processing differences.

What Is the Role of Vowels for Grammar?

Vowels and consonants may play different roles in language acquisition and use. According to Nespor et al. (2003), for example, consonants may be used preferentially for learning and recognizing words, whereas vowels may carry the prosody of utterances; since prosody correlates with syntactic structure (e.g., Nespor & Vogel, 1986), vowels may thus give cues to constituent structure in speech.

How does such a role of the vowels relate to our results? After all, in the simple repetition-based rules we used, participants do not really have to detect constituent structure; they just have to extract relations among vowels. Moreover, repetition-based relations, such as the ones used here, seem to be processed by a specialized "primitive" (Endress et al., 2007); it is thus unclear how such relations mesh with the more abstract syntactic structures for which vowels should be important.

Still, there are linguistic processes that define certain domains and that bear some resemblance to the rules used in our experiments. Vowel harmony is a case in point. In many languages, vowels within a domain have to agree in one or more features. In Hungarian, for instance, vowels within a word have to be either all front or all back (with certain exceptions). Accordingly, suffixes have different forms for front and back vowels; in the dative case, for example, words with back vowels take the *-nak* suffix (such as *sor-nak*, "row"), whereas words with front vowels take the *-nek* suffix (such as *sör-nek*, "beer").

In our experiments, the identity of the first and the last vowel may define domains in a way similar to vowel harmony. Since vowel harmony is much more frequent than consonant harmony in the world's languages (Nespor & Vogel, 1986), such a function of the vowels may explain the advantage we observed. It will thus be important to find out precisely why vowels are favored in our experiments and in the case of harmonic processes, and whether these two phenomena are related.

Can "Structural" Saliency Account for the Vowel Advantage?

Although it is unclear how our results relate to a prosodic function of vowels, they demonstrate that the vowel advantage for generalizations cannot be reduced to their "salience"; in fact, it persists even if vowels are made less salient than consonants. A related possibility, however, is that structural generalizations of the type investigated here are preferentially performed over syllabic *nuclei.*[8] Since the nuclei have to be more sonorant than do either onsets or codas, this possibility would be in some sense a more structural version of the energy hypothesis outlined above.

In our Experiment 3, and in previous experiments (Toro et al., 2008), vowels were always the nuclei of the syllables; likewise, one may consider most of the isolated consonants used in Experiment 4 as essentially nuclear, and results showed some generalization over them. (Similar consonantal nuclei exist in languages such as Croatian and Czech.) Note that the sonorance of the nuclei cannot be equated to their salience or their audibility (see Clements, in press), and, in fact, our results demonstrate that the salience of the vowels is not the crucial determinant of the relative ease of the generalization.

It is thus possible that the structural position within a syllable may constitute an important factor for the relative ease of potential structural generalizations. Such generalizations may be favored in nuclei relative to other positions. Again, it is unclear why the nucleus should be particularly suitable for generalizations, but such a possibility should be explored in further empirical work.

Comparative studies provide further evidence that it is not just how easy it is to recognize an element in a sequence that determines the computations performed over that element; indeed, vowels can be recognized more easily than consonants—for example, in automatic speech recognition

applications (e.g., Benzeghiba et al., 2007). However, in a nonhuman primate (*Saguinus oedipus*) that presumably processes human speech by relying on the acoustic recognizability of speech elements, vowels (rather than consonants, as in humans) are the preferred targets of statistical computations (Newport, Hauser, Spaepen, & Aslin, 2004). Hence, if vowels and consonants were processed differently because vowels are more recognizable, one would expect *all* computations to operate better over vowels than over consonants. However, our observations suggest that structures are easily generalized over vowels, whereas statistical computations are hardly performed over them. This pattern of results, therefore, cannot be due to the relative recognizability of vowels and consonants.

Finally, it is important to note that the relative success in Experiment 4 also has a different interpretation. Indeed, the material in Experiment 4 is also the least "language-like," consisting of a stream of consonant sounds separated by silence. It is thus possible that, for nonlinguistic input, general learning mechanisms enable participants to process regularities over these "consonants." In other words, the failure to generalize rules over consonants might be specific to the role of consonants qua consonants—that is, as constituents of speech. When these sounds are taken out of a linguistic context, as in Experiment 4, they might be processed differently, thereby allowing simple generalizations to be drawn over them. In fact, previous research has shown that the same physical sound, when it is perceived in a linguistic or a nonlinguistic context, is processed differently and induces different patterns of physiological brain responses (Dehaene-Lambertz et al., 2005; Möttönen et al., 2006). Still, the results from the pooled data of Experiments 1 and 2 suggest that some generalizations over consonants are also possible, with more language-like input.

In summary, the present work adds to a growing number of empirical studies that support the hypothesis that consonants and vowels play different roles during language processing (e.g., Bonatti et al., 2005; Caramazza et al., 2000; Cutler et al., 2000; Nazzi, 2005; Owren & Cardillo, 2006; Toro et al., 2008). Importantly, such differences are difficult to account for in terms of lower-level acoustic cues (even though these cues do differentiate vowels and consonants; e.g., Knobel & Caramazza, 2007). There are two interpretations of what might be the source of such differences. They may be due to innate constraints that guide the system to process vowels differently from consonants from very early stages; or acoustical and distributional differences may progressively bias the system toward such differential processing of phonological representations. Be that as it may, the empirical asymmetries between vowels and consonants do exist, and in order to develop a more comprehensive theory of how language is processed, it will be important to understand their origins.

AUTHOR NOTE

This research was funded by McDonnell Foundation Grant 21002089; by CEE Special Targeted Project CALACEI (Contract 12778, NEST); by the Mind, Brain, and Behavior Interfaculty Initiative at Harvard University; and by PRIN2005 to M.N. We thank Jacques Mehler and Luca Bonatti for their valuable insights, and Erika Marchetto for lending us her voice for Experiment 4. Correspondence concerning this article should be addressed to A. D. Endress, Harvard University, 1052 William James Hall, 33 Kirkland Street, Cambridge, MA 02138 (e-mail: ansgar.endress@m4x.org).

Note—Accepted by the editorial board of Editor-Elect Jeremy M. Wolfe.

REFERENCES

Aslin, R., Saffran, J., & Newport, L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, **9**, 321-324.

Benzeghiba, M., De Mori, R., Deroo, O., Dupont, S., Erbes, T., Jouvet, D., et al. (2007). Automatic speech recognition and speech variability: A review. *Speech Communication*, **49**, 763-786.

Bonatti, L. L., Peña, M., Nespor, M., & Mehler, J. (2005). Linguistic constraints on statistical computations: The role of consonants and vowels in continuous speech processing. *Psychological Science*, **16**, 451-459.

Bonatti, L. L., Peña, M., Nespor, M., & Mehler, J. (2007). On consonants, vowels, chickens, and eggs. *Psychological Science*, **18**, 924-925.

Caramazza, A., Chialant, D., Capasso, R., & Miceli, G. (2000). Separable processing of consonants and vowels. *Nature*, **403**, 428-430.

Chomsky, N. (1980). *Rules and representations*. Oxford: Blackwell.

Christophe, A., Nespor, M., Guasti, M. T., & Van Ooyen, B. (2003). Prosodic structure and syntactic acquisition: The case of the head-direction parameter. *Developmental Science*, **6**, 211-220.

Clements, N. (in press). Does sonority have a phonetic basis? Comments on the chapter by Vaux. In E. Raimy & C. Cairns (Eds.), *Contemporary views on architecture and representations in phonological theory*. Cambridge, MA: MIT Press.

Cutler, A., Sebastián-Gallés, N., Soler-Vilageliu, O., & van Ooijen, B. (2000). Constraints of vowels and consonants on lexical selection: Cross-linguistic comparisons. *Memory & Cognition*, **28**, 746-755.

Dehaene-Lambertz, G., Pallier, C., Serniclaes, W., Sprenger-Charolles, L., Jobert, A., & Dehaene, S. (2005). Neural correlates of switching from auditory to speech perception. *NeuroImage*, **24**, 21-33.

Dutoit, T., Pagel, V., Pierret, N., Bataille, F., & vander Vrecken, O. (1996, October). *The MBROLA project: Towards a set of high-quality speech synthesizers free of use for non-commercial purposes*. Paper presented at the proceedings of the Fourth International Conference on Spoken Language Processing, Philadelphia.

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development.* Cambridge, MA: MIT Press.

Endress, A., & Bonatti, L. (2007). Rapid learning of syllable classes from a perceptually continuous speech stream. *Cognition*, **105**, 247-299.

Endress, A., Dehaene-Lambertz, G., & Mehler, J. (2007). Perceptual constraints and the learnability of simple grammars. *Cognition*, **105**, 577-614.

Endress, A., Scholl, B., & Mehler, J. (2005). The role of salience in the extraction of algebraic rules. *Journal of Experimental Psychology: General*, **134**, 406-419.

Garcia, J., Hankins, W. G., & Rusiniak, K. W. (1974). Behavioral regulation of the milieu interne in man and rat. *Science*, **185**, 824-831.

Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition*, **73**, 135-176.

Gleitman, L. R., & Wanner, E. (1982). Language acquisition: The state of the state of the art. In E. Wanner & L. R. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 3-48). Cambridge: Cambridge University Press.

Gómez, R., Gerken, L., & Schvaneveldt, R. (2000). The basis of transfer in artificial grammar learning. *Memory & Cognition*, **28**, 253-263.

Keidel, J. S., Jenison, R. L., Kluender, K. R., & Seidenberg, M. S. (2007). Does grammar constrain statistical learning? Commentary on Bonatti et al. (2005). *Psychological Science*, **18**, 922-923.

KNOBEL, M., & CARAMAZZA, A. (2007). Evaluating computational models in cognitive neuropsychology: The case from the consonant/vowel distinction. *Brain & Language*, **100**, 95-100.
LADEFOGED, P. (2001). *Vowels and consonants: An introduction to the sounds of language.* Oxford: Blackwell.
LADEFOGED, P. (2006). *A course in phonetics* (5th ed.). Boston: Wadsworth.
MANI, N., & PLUNKETT, K. (2007). Phonological specificity of vowels and consonants in early lexical representations. *Journal of Memory & Language*, **57**, 252-272.
MCCARTHY, J. J. (1979). *Formal problems in Semitic phonology and morphology*. Doctoral dissertation, MIT, Cambridge, MA.
MCCLELLAND, J. L., RUMELHART, D. E., & THE PDP RESEARCH GROUP (EDS.) (1986). *Parallel distributed processing, Vol. 2: Psychological and biological models*. Cambridge, MA: MIT Press.
MEHLER, J., DUPOUX, E., NAZZI, T., & DEHAENE-LAMBERTZ, G. (1996). Coping with linguistic diversity: The infant's viewpoint. In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax* (pp. 101-116). Mahwah, NJ: Erlbaum.
MEHLER, J., PEÑA, M., NESPOR, M., & BONATTI, L. L. (2006). The "soul" of language does not use statistics: Reflections on vowels and consonants. *Cortex*, **42**, 846-854.
MONAGHAN, P., & SHILLCOCK, R. (2003). Connectionist modeling of the separable processing of consonants and vowels. *Brain & Language*, **86**, 83-98.
MONAGHAN, P., & SHILLCOCK, R. (2007). Levels of description in consonant/vowel processing: Reply to Knobel and Caramazza. *Brain & Language*, **100**, 101-108.
MORGAN, J. L., & DEMUTH, K. (1996). *Signal to syntax: Bootstrapping from speech to grammar in early acquisition.* Mahwah, NJ: Erlbaum.
MÖTTÖNEN, R., CALVERT, G., JÄÄSKELÄINEN, I., MATTHEWS, P., THESEN, T., TUOMAINEN, J., & SAMS, M. (2006). Perceiving identical sounds as speech or non-speech modulates activity in the left posterior superior temporal sulcus. *NeuroImage*, **30**, 563-569.
NAZZI, T. (2005). Use of phonetic specificity during the acquisition of new words: Differences between consonants and vowels. *Cognition*, **98**, 13-30.
NAZZI, T., & NEW, B. (2007). Beyond stop consonants: Consonantal specificity in early lexical acquisition. *Cognitive Development*, **22**, 271-279.
NESPOR, M., PEÑA, M., & MEHLER, J. (2003). On the different roles of vowels and consonants in speech processing and language acquisition. *Lingue & Linguaggio*, **2**, 201-227.
NESPOR, M., & VOGEL, I. (1986). *Prosodic phonology.* Dordrecht, The Netherlands: Foris.
NEWPORT, E. L., & ASLIN, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, **48**, 127-162.
NEWPORT, E. L., HAUSER, M. D., SPAEPEN, G., & ASLIN, R. N. (2004). Learning at a distance II. Statistical learning of non-adjacent dependencies in a non-human primate. *Cognitive Psychology*, **49**, 85-117.
OWREN, M. J., & CARDILLO, G. C. (2006). The relative roles of vowels and consonants in discriminating talker identity versus word meaning. *Journal of the Acoustical Society of America*, **119**, 1727-1739.
PINKER, S. (1999). *Words and rules: The ingredients of language.* New York: Harper Collins.
SAFFRAN, J., ASLIN, R., & NEWPORT, E. (1996). Statistical learning by 8-month-old infants. *Science*, **274**, 1926-1928.
SELKIRK, E. (1974). French liaison and the $\overline{X}$ notation. *Linguistic Inquiry*, **5**, 573-590.
SHARP, D., SCOTT, S., CUTLER, A., & WISE, R. (2005). Lexical retrieval constrained by sound structure: The role of the left inferior frontal gyrus. *Brain & Language*, **92**, 309-319.
TOMASELLO, M. (2000). The item-based nature of children's early syntactic development. *Trends in Cognitive Sciences*, **4**, 156-163.
TORO, J. M., NESPOR, M., MEHLER, J., & BONATTI, L. (2008). Finding words and rules in a speech stream: Functional differences between vowels and consonants. *Psychological Science*, **19**, 137-144.
TUNNEY, R., & ALTMANN, G. (2001). Two modes of transfer in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 614-639.
VAN OOIJEN, B. (1996). Vowel mutability and lexical selection in English: Evidence from a word reconstruction task. *Memory & Cognition*, **24**, 573-583.

NOTES

1. In Newport and Aslin's (2004) experiments, vowel sequences could be repeated in adjacent words; this manipulation apparently makes vowel sequences pop out, because Bonatti et al. (2005) also observed a sensitivity to nonadjacent vowel relations under these conditions. When these repetitions were removed, however, participants used only consonants, not vowels.

2. Vowels also may carry other information, such as cues to the identity of the talker (Owren & Cardillo, 2006), but these functions are not relevant to the present experiments.

3. TPs are conditional probabilities of encountering a syllable after having encountered another syllable. Conditional probabilities like $P(\sigma_{i+1} = /\text{key}/ \mid \sigma_i = /\text{don}/)$ are high within words and low between words (σ denotes syllables in a speech stream).

4. Technically, in CVCVCV words, the mutual information between the consonant frames and the words is higher than that between the vowel frames and the words.

5. No direct comparisons between performances in both tasks are included, since we have no independent motivation that would predict their relative performance. For example, some rule-like generalizations are affected differently from statistical processes by different manipulations on the familiarization material (e.g., Endress & Bonatti, 2007). Although the generalizations studied here are arguably rather different from those studied by Endress and Bonatti, we do not know how different parameters of the familiarization stream affect the performance on the two tasks, and we cannot be sure whether the relative performance on the two tasks would be representative of other parameter sets. We thus believe that the use of inferential statistics would not be justified due to these concerns, but that, in any case, the relative performance on the two tasks is not crucial to our conclusions.

6. For the exact algorithm used, see www.fon.hum.uva.nl/praat/manual/Sound__Get_energy___.html.

7. The results from the first 15 participants did not reach significance ($M = 58.3\%$, $SD = 18.7\%$) [$t(14) = 1.72$, $p = .106$, n.s.].

8. In phonological theory, a syllable is a hierarchical structure, starting with the *onset*—that is, the initial consonant (cluster)—which is followed by the *rhyme*; the rhyme is composed of the nucleus (generally a vowel), followed by a *coda*—that is, the final consonant (cluster). (The only obligatory part is the nucleus; onset and coda are optional in all languages.) In the word /test/, for example, /t/ is the onset, /e/ is the nucleus, and /st/ is the coda; the syllable thus has the structure [t[[e][st]]].

(Manuscript received March 19, 2008;
revision accepted for publication June 8, 2008.)