Language universals in human brains

Iris Berent†, Tracy Lennertz*, Jongho Jun‡, Miguel A. Moreno§, and Paul Smolensky

Do speakers know universal restrictions on linguistic elements that are absent from their language? We report an experimental test of this question. Our case study concerns the universal restrictions on initial consonant sequences, onset clusters (e.g., b/l in block). Across languages, certain onset clusters (e.g., lb) are dispreferred (e.g., systematically under-represented) relative to others (e.g., bh). We demonstrate such preferences among Korean speakers, whose language lacks initial C1C2 clusters altogether. Our demonstration exploits speakers’ well known tendency to misperceive ill-formed clusters. We show that universally dispreferred onset clusters are more frequently misperceived than universally preferred ones, indicating that Korean speakers consider the former cluster-type more ill-formed. The misperception of universally ill-formed clusters is unlikely to be due to a simple auditory failure. Likewise, the aversion of universally dispreferred onset clusters by Korean speakers is not explained by English proficiency or by several phonetic and phonological properties of Korean. We conclude that language universals are neither relics of language change nor are they artifacts of generic limitations on auditory perception and motor control—they reflect universal linguistic knowledge, active in speakers’ brains.

The “nature vs. nurture” debate concerns the origin of speakers’ knowledge of their language. Both sides of this controversy presuppose that people have some knowledge of abstract linguistic regularities. They disagree on whether such regularities reflect the properties of linguistic experience, auditory perception, and motor control (1, 2) or universal, possibly innate, and domain-specific restrictions on language structure (3–5, **). Empirical support for such restrictions comes from linguistic universals: regularities exhibited across the world’s languages. These universals, for example, assert that the sound sequence b/lbf makes a poor word, whereas the sequence b/lbf is better: Languages always make use of words like b/lbf before (as in Russian) resorting to words like b/lbf. But the significance of such observations is unclear. One view holds that language universals form part of the language faculty of all speakers (5–7). The alternative denies that speakers have knowledge of language universals. Rather, speakers simply know regularities (either structural or statistical) concerning words in their own language. Language universals are not mentally represented—they are only statistical tendencies, shaped by generic (auditory and motor) constraints on language evolution (8). For example, words beginning with lb have a tendency to decline relative to those beginning with bl because the former are more frequently mispronounced or misperceived. The question at hand, then, is whether language universals are active in the brains of all speakers, or more relics of systematic language change and its distal generic causes?

The matter is difficult to resolve because it is not easy to distinguish active knowledge of a universal regularity from “mere analogy” with memorized expressions that happen to exhibit that regularity. English speakers showing preferences for syllables like b/lf over l/bf might be demonstrating knowledge of the relevant general universals, or they might be reflecting only their knowledge that English has words relevantly like b/lf (e.g., blip) but not l/bf. The strategy we use to distinguish these possibilities exploits universal preferences among types of words, all of which are absent from a speaker’s language (see also refs. 9–12). Here, we examine whether the universal preference for syllables like b/lf over l/bf is available to Korean speakers, whose language arguably gives them no experience with words beginning with two consonants (13–18, ‡‡). Such knowledge may be either genetically predetermined (3, 19) or partly learned based on properties of the human speech production, perception, and cognition system (20); the work discussed here does not speak to this question.

To probe for universal linguistic knowledge, we compare the preferences of Korean speakers for a scale of four types of syllable-initial consonant sequences (“onset clusters”), such as the C1C2 sequence in C1C2VC1 words (C and V denote consonant and vowel, respectively). At the top of the scale are clusters as in b/lbf, at the bottom are clusters as in lb/lbf, and intermediate are those in bn/lbf and b/lbf. An analysis of a diverse language sample (data from ref. 21 reanalyzed in ref. 22) shows that the frequency of such onset clusters decreases monotonically and reliably across the hierarchy (e.g., the bd-type is more frequent than the lb-type). Moreover, languages allowing less frequent (e.g., lb-type) clusters tend to allow more frequent ones (e.g., bd-type). Such preferences have been attributed to the abstract property of sonority (§)—approximately correlated with the physical energy of speech sounds—which is key to many universals concerning the arrangement of speech sounds in words and syllables. Least sonorous, with $s = 1$, are stops, such as p, t, k, b, d, and g, and fricatives, such as f and v, next, with $s = 2$, are nasals, such as n and m; then, liquids l and r with $s = 3$; and, finally, glides $w$ and $y$, with $s = 4$. Accordingly, the C1C2 cluster in b/lbf manifests a large rise in sonority ($s = 1$) b/lbf, manifests a smaller rise ($s = 1$), b/lbf exhibits a plateau ($s = 0$), and l/bf manifests a sonority fall ($s = 2$). Crucially, the larger the sonority distance $s$ the more preferred the syllable across languages. The universal preference (denoted by a sign) is thus Large Rises Small Rises Plateau Falls (e.g., bl bn bd lb) (23, 24). Our question is whether speakers of all languages exhibit active knowledge of this scale.

Author contributions: I.B., T.L., and P.S. designed research; M.A.M. performed research; I.B. and J.J. analyzed data; and I.B. and P.S. wrote the paper.

The authors declare no conflict of interest.

FREELY AVAILABLE ONLINE THROUGH THE PNAS OPEN ACCESS OPTION.

‡‡We assume such restrictions form part of the grammar—a computational faculty of the brain and mind that can generate an infinite number of sentences from a finite set of operations on linguistic variables. The existence of language universals may thus reflect the presence of active universal constraints in the grammars of all speakers.

© 2008 by The National Academy of Sciences of the USA
According to one linguistic theory, optimality theory (5, 7, 25), cross-linguistic generalizations arise from universal knowledge, active in the language faculty of all speakers irrespective of the actual words in their language. Accordingly, we expect that speakers will favor blif ( < 0) to lbif ( > 2) even though their language may have neither type of syllable (e.g., English) or indeed no CCVC syllables at all (e.g., Korean).

To test for active knowledge of this universal scale, we capitalized on the following principle: When presented with a speech sound sequence that is ill-formed in their language, listeners tend to repair it in perception as a better-formed one. For example, given an illicit consonant sequence (e.g., /tla/), speakers misperceive the problematic sequence as separated by a short schwa-like vowel that we write as “e” [e.g., / tela/ (26, 27)]. Extending this principle, we predict that if speakers actively deploy knowledge of universal principles, then universally less-preferred clusters should be more likely to be misperceived compared with universally more-preferred ones. And indeed English speakers are most likely to misperceive highly ill-formed clusters like /lbif/ (as /lbef/). They are somewhat less likely to misperceive /bdlf/ and still less likely to misperceive /bntf/ (22).

Such misperception is not due to a simple failure to encode the acoustic properties of the initial consonants—participants are demonstrably able to represent dispreferred consonant sequences accurately when attention to phonetic information is encouraged (see ref. 22, experiments 5 and 6). This evidence is thus consistent with the hypothesis that English speakers actively deploy knowledge of sonority sequencing universals.15 It is quite unclear how mere knowledge of English words could explain such misperceptions, given that no English syllable begins in /b/, /bd/, or /bn/. It is imaginable, however, that the observed pattern might be explained by some sort of generalization mechanism—as yet unspecified—operating solely on the existing consonant–consonant (CC)-syllable initial sequences of English words.

Thus, we turn to native speakers of Korean, whose language provides no CC-initial syllables. If these speakers’ patterns of perceiving CCVC words follows the sonority sequencing universals, then it seems virtually impossible that this behavior reflects some (unspecified) generalization from Korean words alone.

We tested Korean participants’ tendency to misperceive CCVC words as CeCVC in two-ways. Experiment 1 exploits the difference in syllable count between monosyllabic CCVC and their “repair,” CeCVC. In each trial, participants were presented with an auditory item, either a C1C2VC3 nonword (e.g., /lbif/) or its disyllabic counterpart C1eC2VC3 (e.g., /lbef/). Each CCVC item was classified as a sonority Fall, Plateau, Small Rise, or Large Rise. Participants were asked to indicate whether the stimulus included one syllable or two by pressing the corresponding computer key (a syllable count task). Because Korean frequently repairs loanwords, including illicit CC sequences by inserting a schwa-like vowel (14), we expected participants to misperceive the C1C2VC3 forms as C1eC2VC3 (e.g., /lbif 3 lbef/). Of interest is whether such perceptual repair depends on the universal well-formedness of the C1C2 sequence, determined by the sonority distance $s$ $\epsilon$ (C2) $\notin$ (C1). In a second experiment, we used a more direct measure of participants’ tendency to misperceive illicit clusters: We simply asked them, “is /lbif/ identical to /lbef/?” Participants were presented with two auditory stimuli—either identical (e.g., /lbif-lbif, lbef-lbef/) or repair related (e.g., /lbif-lbef/ and asked to determine whether the two stimuli were identical (an identity judgment task). If the universal preference for greater sonority distance is active in participants’ linguistic knowledge, they should more often misjudge C1C2VC3 for C1eC2VC3 if sonority distance is smaller.

Results

Experiment 1: Syllable Count. In Experiment 1, sonority distance affected the perception of both monosyllabic items and their disyllabic counterparts. The 2 (syllable) $\times$ 4 (onset-type) ANOVAs conducted by using participants (F1) and items (F2) as random variables each yielded a significant interaction [accuracy: F1 (3, 54) 10.10, P < 0.0001; F2 (3, 87) 34.91, P < 0.0001; response time: F1 (3, 51) 5.78, P = 0.0002; F2 (3, 84) 7.79, P < 0.0001].

A test of the simple main effect of onset-type showed that sonority distance significantly modulated the perception of monosyllabic items [accuracy: F1 (3, 54) 8.87, P < 0.0002; F2 (3, 87) 17.3, P < 0.0001; response time: F1 (3, 54) 4.36, P = 0.0099; F2 (3, 87) 6.77, P = 0.0005]. As sonority distance (universal well-formedness) decreased, responses to monosyllabic items were slower and less accurate, indicating that participants tended to misperceive them as disyllabic (See Fig. 1). Planned comparisons of responses to monosyllabic items showed that participants were more accurate and significantly faster responding to onsets with large sonority rises compared with relatively less well formed onsets with plateaus [accuracy: F1

---

15Regression analyses yielded a significant effect of onset-type after controlling for various measures of consonant coocurrence in English.
(54) 1.92, P 0.07; F2 (87) 2.67, P 0.01; response time: F1 (54) 2.68, P 0.01; F2 (87) 2.79, P 0.007), which, in turn, yielded significantly more accurate responses relative to still less well formed onsets of falling sonority (F1 (54) 2.95, P 0.005; F2 (87) 4.12, P 0.0001; response time, P 0.12, n.s.). Responses to onsets with large and small rises did not differ reliably (P 0.05). Thus, Korean speakers tend to misperceive universally ill-formed onsets as disyllabic, a result consistent with findings on English speakers (22).

Remarkably, the universal well-formedness of monosyllabic onsets also modulated responses to their disyllabic counterparts. Disyllabic counterparts of well formed monosyllables (e.g., beltif, counterpart of blif) produced slower and less accurate responses than disyllabic counterparts of less well formed monosyllables (e.g., lebitif, counterpart of lbif). The simple main effect of onset-type was significant [for accuracy F1 (3, 54) 6.80, P 0.0007; F2 (3, 87) 16.83, P 0.0002; for response time, F1 (3, 51) 2.48, P 0.08; F2 (3, 84) 1.79, P 0.16]. Planned comparisons showed that disyllabic counterparts of more well formed onsets with large rises produced slower responses than counterparts of plateaus (response time: F1 (54) 1.93, P 0.06; F2 (84) 2.18, P 0.04; accuracy: F1 (54) 1.46, P 0.16; F2 (87) 2.3, P 0.03), which, in turn, produced significantly less accurate responses than the counterparts of the least well formed onsets with sonority falls [accuracy: F1 (54) 2.84, P 0.007; F2 (84) 4.47, P 0.0001; response time: both P 1]. Responses to onsets with large and small sonority rises did not differ significantly (all P values 0.05). Auxiliary stepwise regression analysis demonstrated that the difficulty of processing CeCVC disyllables was not due to the phonetic length of the vowel itself, because the effect of onset-type remained significant after controlling for vowel length in the first step [R change 0.244, F2 (1, 117) 38.19, P 0.0001]. This effect, observed also with English speakers, might be due to competition between faithfully perceived and misperceived forms: Participants are less likely to misperceive CeCVC3 as CeCVC1 when the sequence CeC3 is more ill-formed. Thus, the universal ill-formedness of lbif protects lebitif from such errors.

**Experiment 2: Identity Judgment.** The tendency of Korean speakers to consider universally ill-formed C1C2VC3 onsets as disyllabic suggests that such onsets are repaired in perception as C1C2VC3 sequences. If this interpretation is correct, then participants should also consider ill-formed C1C2VC3 onsets as identical to their C1C2VC3 counterparts. The results of the identity-judgment task (see Fig. 2) are consistent with this prediction. A one-way ANOVA indicated that, as sonority distance (well-formedness) decreased, people were significantly less accurate [F1 (3, 135) 104.39, P 0.0002; F2 (3, 87) 22.63, P 0.0002] and significantly slower [F1 (3, 114) 74.03, P 0.0002; F2 (3, 84) 12.94, P 0.0002] in determining that monosyllabic items C1C2VC3 were nonidentical to their disyllabic counterparts C1eC2VC3 (e.g., blitif–lebitif). Planned comparisons showed that participants were significantly more accurate and significantly faster in responding to pairs whose monosyllabic member had well formed onsets with a large sonority rise (e.g., beltif) compared with less well formed onsets with plateaus [accuracy: F1 (135) 10.30, P 0.0001; F2 (87) 5.07, P 0.0001; response time: F1 (114) 4.01, P 0.0002; F2 (84) 2.80, P 0.007], which, in turn, yielded significantly more accurate and significantly faster responses compared with ill-formed onsets with sonority falls [accuracy: F1 (135) 5.12, P 0.0001; F2 (87) 2.00, P 0.05; response time: F1 (114) 8.91, P 0.0001; F2 (84) 1.93, P 0.06]. Responses to onsets with large and small sonority rises did not differ significantly (P 0.05). Thus, Korean speakers misperceive universally ill-formed onsets with small sonority distances as identical to their disyllabic counterparts.

**Discussion**

The results of Experiments 1 and 2 reveal a striking correspondence between the behavior of Korean speakers and the distribution of initial CC sequences across languages. Across languages, initial C1C2 sequences are less frequent the smaller the sonority distance, and languages allowing sequences with a given sonority distance also allow sequences with greater distances. The experimental results show that initial C1C2 sequences with smaller sonority distances are systematically misperceived: As sonority distance decreases, monosyllabic C1C2VC3 items are more frequently judged disyllabic (in Experiment 1) and more often considered identical to their disyllabic counterpart C1eC2VC3 (in Experiment 2).

The difficulty of distinguishing low sonority distance sequences from their disyllabic counterparts is not due to a simple auditory failure, because accuracy with identical items was nearly perfect (M 0.98, for both monosyllabic and disyllabic items). Moreover, the effect of universal C1C2VC3 ill-formedness extended even to disyllabic C1eC2VC3 forms (in Experiment 1): People were more accurate responding to the disyllabic counterparts of ill-formed sequences (e.g., blitif, counterpart of lbif) relative to counterparts of well formed ones (e.g., beltif, counterpart of blif). The persistent aversion to these ill-formed sequences, even when they are not physically present, cannot stem from difficulty in their auditory perception. It is also unlikely to result from motor difficulties in their pronunciation, because participants did not articulate the sequences overtly. Even if participants somehow engaged in covert articulation (a possibility for which we have no evidence), it is hard to see why they should experience difficulties with forms like beltif compared with lebitif.

The preferences of Korean speakers are also not explained by their proficiency with second languages that allow initial CC sequences, such as English. Although most participants had some level of English proficiency, this factor did not modulate the effect of onset-type in either Experiment 1 or 2 (see SI). Moreover, the sensitivity of Korean participants in the identity judgment task (d 2.15) was even higher than native English speakers (d 1.82) described in ref. 22—a finding that clearly counters the attribution of sonority distance effects to English proficiency. Likewise, the results are not due to several conceivably relevant phonological and phonetic characteristics particular to Korean, including the phonetic release of initial stop-consonants, their voicing, the distribu-
Fig. 3. The correlation between the distribution of onset-cluster types across languages and the performance of Korean and English speakers. The abscissa provides the frequency of four onset types in a sample of 90 languages (21, 22). The ordinate reflects the accuracy of Korean and English speakers (English data from ref. 22) to such onsets in syllable count and identity judgment tasks.

tion of [l] and [l] allophones, the experience with Korean words beginning with consonant-glide sequences, and the occurrence of the CC sequence across Korean syllables (see 25).

Despite little or no experience with initial consonant sequences in their language, Korean speakers demonstrate preferences concerning such sequences, preferences that mirror the distribution of these sequences across languages and that converge with those of English speakers (see Fig. 3). This convergence is all of the more remarkable in view of the linguistic differences between these languages. The results are consistent with the hypothesis that adult human brains possess knowledge of universal properties of linguistic structures absent from their language.

Several limitations of our conclusions are noteworthy. In view of the confinement of our present investigation to a handful of languages, a full evaluation of the universality of the sonority hierarchy requires extensions to additional languages that further restrict the occurrence of consonant clusters. Our results also cannot determine the origins of the sonority hierarchy. Although the behavior of participants in our experiments reflects abstract phonological knowledge and not merely difficulties in the perception and articulation of such clusters, such knowledge is probably not arbitrary. A growing body of research has suggested that phonological restrictions in general and the restrictions on sonority in particular might be grounded in the phonetic properties of speech perception and articulation (20, 28, 29). The role of such phonetic pressures in shaping speakers’ abstract phonological knowledge concerning sonority awaits further research. Finally, the hypothesis that speakers possess universal phonological constraints does not imply that knowledge of those constraints is experience-independent nor does it speak to its species-specificity and evolutionary origins (30, 31). How speakers of different languages converge on the same universal knowledge remains to be seen.

**Methods**

**Experiment 1. Participants.** Participants were 19 native-Korean speakers, students at Gyeongsang National University in South Korea.

**Materials.** The materials consisted of 120 pairs of monosyllabic nonwords and their disyllabic counterparts, used with English-speaking participants in ref. 22, experiments 1–4. Monosyllabic words were CVCVC words arranged in 30 quartets (see 31). Most quartet members (113 of 120 items) shared their rhyme and differed on the structure of their CVC clusters. Onset clusters were of four types. One type had a large sonority rise (e.g., bigil); in a second type (25 of 30 members had a smaller rise (e.g., bitil) 31 a third category had a sonority plateau (e.g., bitil) and the final category had a fall in sonority (e.g., bitil). Stimuli were produced by a speaker of Russian, in which all types of onset clusters used are attested. (A speaker of another language might have introduced a serious artifact by producing less fluently those onset cluster types not attested in their native language.)

**Procedure.** Participants were seated near a computer and wore headphones. Participants initiated each trial by pressing the space bar. In a trial, they were presented with an auditory stimulus. Participants were asked to indicate as quickly and accurately as possible whether the item had one syllable or two by pressing one of two keys (1 one syllable; 2 two syllables). Immediately before the experimental session, participants were presented with a practice phase. Because it was (by design) impossible to illustrate the task with Korean words, we encouraged participants with 14 practice items in English (e.g., sport vs. support) and provided feedback on their accuracy (“correct” and “incorrect” responses). Outliers (correct responses falling 2.5 SD beyond the mean, 3% of the data) were excluded from the analyses of response time in Experiments 1 and 2. Response times are reported from the onset of the auditory stimulus.

**Experiment 2. Participants.** Participants were 46 native-Korean speakers, students at Gyeongsang National University in South Korea.

**Materials.** The materials consisted of the same items from Experiment 1. They were arranged in pairs: Half were identical (either monosyllabic or disyllabic), and half repair-related (e.g., bitil-bigil or belit-bitil). The materials were arranged in two lists, matched for the number of stimuli per condition (onset type: identity order) and counterbalanced such that, within a list, each item appeared in either the identity or the nonidentity condition but not both. Each participant was assigned to one list.

**Procedure.** Participants were seated near a computer wearing headphones. Participants initiated each trial by pressing the space bar. In a trial, they were presented with two auditory stimuli (with an onset asynchrony of 1,200 ms) and they were asked to indicate whether the two stimuli were identical by pressing the 1 or 2 keys for “identical” and “nonidentical” responses, respectively. Slow responses (response time > 3,500 ms) received a computerized warning signal. Before the experiment, participants were given a short practice session. As in Experiment 1, it was impossible to illustrate the task using Korean words, and consequently, we used English examples (e.g., plight-plight vs. polite-plight). During the practice session, participants received computerized feedback for both accuracy and speed. Response times are reported from the onset of the second auditory stimulus.

**ACKNOWLEDGMENTS.** We thank Roger Shepard and three anonymous reviewers for comments and Yang Lee for facilitating the testing of participants at Gyeongsang National University. This work was supported by National Institute of Deafness and Other Communication Disorders Grant R01 DC003277 (to J.B.) and National Institute of Child Health and Human Development Grant HD-01994 (to Haskins Laboratories).

---