

# CFA/VISHNO 2016

## Quelle adaptation acoustique pour les langues du monde ?

C. Coupe<sup>a</sup> et I. Maddieson<sup>b</sup>

<sup>a</sup>Laboratoire Dynamique du Langage, Institut des Sciences de l'Homme, 14 Avenue Berthelot, 69007 Lyon, France

<sup>b</sup>Department of Linguistics, University of New Mexico, MSC03 2130, 1 University of New Mexico, Albuquerque, 87131-0001, USA  
christophe.coupe@cns.fr



LE MANS

L'hypothèse d'adaptation acoustique postule que les espèces utilisent des signaux de communication dont les caractéristiques optimisent la transmission d'information dans l'environnement qu'elles occupent. Si ce phénomène est attesté chez différentes espèces animales, la question se pose si la grande diversité sonore des langues du monde reflète en partie des adaptations différenciées à différents environnements de vie. Nous défendons l'idée que la densité de végétation exerce sur les langues du monde une pression à l'encontre des consonnes, en particulier obstruantes, et des syllabes de structure complexe. Nous proposons un mécanisme causal, à savoir qu'une végétation dense dégrade préférentiellement les ondes sonores de haute fréquence sur lesquelles reposent la perception et l'identification des éléments précédents. Nous nous appuyons pour cela sur une étude géostatistique des caractéristiques sonores de près de 700 langues du monde.

## 1 Cadre théorique et pratique

### 1.1 La notion d'adaptation acoustique

De nombreuses expériences démontrent l'impact d'un environnement de propagation sur la structure acoustique d'un son [1,2,3,4]. Sur cette base, l'hypothèse d'adaptation acoustique postule que les espèces utilisent des signaux de communication dont les caractéristiques optimisent la transmission d'information dans l'environnement qu'elles occupent, en particulier pour communiquer à longue distance [5,6]. Cette adaptation serait le produit de la sélection naturelle, de la même façon que pour d'autres aspects physiologiques et cognitifs.

L'adaptation acoustique est étudiée depuis plusieurs décennies par les éthologues [7]. Les oiseaux ont fait l'objet de nombreuses études, mais également les anoues et les mammifères [8,9]. Les conclusions de ces travaux sont que l'adaptation acoustique constitue effectivement l'une des forces évolutives qui expliquent la structure des signaux de communication animaux. Ses effets ne sont toutefois pas toujours clairement observables, car différentes contraintes pèsent sur les signaux de communication et ne vont pas toutes dans le sens d'une maximisation de la distance de propagation des signaux : capacités physiologiques d'émission et de réception des signaux, phylogénie, pressions de prédation, pressions sociales telles le « parasitage » de l'information, système d'appariement, ou encore écologie de l'espèce. Chez certaines espèces d'oiseaux, il a ainsi été montré que des informations sont transmises de façon différenciée selon la distance : par exemple, transmission à longue distance de l'identité de l'espèce mais transmission à plus faible distance de l'identité de l'émetteur [10].

L'adaptation acoustique peut être envisagée en termes de différences interspécifiques, mais également de différences intra-spécifiques. Dans le second cas, c'est la plasticité du système de production sonore d'une même espèce qui est mobilisée. Des études illustrent par exemple comment des espèces d'oiseaux augmentent les fréquences des notes de leur chant dans des environnements où le bruit ambiant est élevé [11,12].

### 1.2 Le cas de l'espèce humaine

Les quelques 6000 langues du monde soulignent la très grande variabilité intra-spécifique de la communication humaine. Cette diversité se manifeste entre autres au niveau des inventaires de sons, et touche des populations qui vivent dans des environnements variés : zones désertiques torrides ou glacées, forêts équatoriales denses et humides, zones tempérées à la végétation plus ouverte etc. Cette situation amène à se poser la question : une partie de cette diversité peut-elle s'expliquer par des adaptations

différenciées aux différents environnements dans lesquels évoluent les êtres humains ?

Il convient pour aborder cette question de tout d'abord distinguer les possibles influences directes et indirectes de l'environnement – au sens large : climat, topographie, végétation etc. – sur les structures linguistiques. Les influences indirectes correspondent à des conditions écologiques qui induisent des situations sociolinguistiques qui à leur tour façonnent les formes linguistiques. Un exemple est celui de Nettle [13] en Afrique subsaharienne : plus le risque écologique dans une région est forte, plus les habitants, qu'ils soient agriculteurs ou éleveurs, font face à des risques élevés de subsistance. Ceci les amène à développer des réseaux d'échanges réciproques sur des aires géographiques assez importantes pour s'assurer mutuellement en cas de crise locale. Cette structuration sociogéographique a pour conséquence d'amoinrir la diversité linguistique et d'observer des langues plus étendues spatialement et au nombre de locuteurs plus important. Un autre exemple est que les climats chauds encouragent une vie en extérieur et donc une communication plus à distance, qui bénéficie de l'emploi de syllabes CV et de consonnes sonantes (produites avec un flux d'air non obstrué dans le tractus vocal, comme [m] ou [l]) [14]. Dans de telles situations, les changements linguistiques observés sont bien causés par l'environnement, mais cette causation est de « second ordre » si on la rapporte à des effets plus directs de l'environnement comme dans le cas de l'adaptation acoustique, qui opère sans médiation des structures populationnelles [15].

Plus directement, les propriétés physiques de l'environnement se manifestent tout d'abord dans la façon dont différentes langues s'appuient sur elles pour construire leur sémantique. L'encodage linguistique de l'espace et des directions peut reposer sur un système autonome comme celui des points cardinaux, mais il peut aussi s'appuyer sur des éléments comme la direction des vents dominants [16] ou le rapprochement / l'éloignement par rapport à un lieu saillant du paysage [17].

L'hypothèse d'adaptation acoustique délaisse quant à elle le domaine de la signification pour se porter sur la nature acoustique des productions linguistiques. Des études publiées ces dernières années ont mis en avant différentes adaptations possibles. Everett [18] a ainsi proposé qu'une altitude élevée favorise la production des consonnes dites éjectives, qui nécessitent une compression de l'air dans la cavité laryngale. L'effort à fournir pour cette compression est en effet atténué lorsque la pression atmosphérique est plus faible, comme cela est le cas en altitude. Une étude de la distribution des consonnes éjectives à la surface du globe suggèrent que celles-ci se rencontrent préférentiellement à des altitudes importantes. De leur côté, Everett, Blasi, & Roberts [19] expliquent le fait que les langues tonales sont

rares dans les régions où l'aridité est importante. Sur la base d'études en laryngologie, ils défendent l'idée qu'un air très aride assèche partiellement les cordes vocales, ce qui rend le contrôle de leurs vibrations plus délicats. Une maîtrise amoindrie de la fréquence fondamentale de la voix rendrait alors difficile l'émergence ou la préservation de distinctions tonales. Dans ces deux exemples, les mécanismes causaux « de premier ordre » mis en exergue s'exercent au niveau de la production de la parole, mais une adaptation acoustique peut également s'opérer face aux conditions de propagation des signaux, comme proposé dans notre étude.

### 1.3 Méthodologie de recherche

Mettre en lumière de façon tangible un phénomène d'adaptation acoustique au sein des langues humaines repose souvent sur la proposition d'un mécanisme causal plausible, associé à une étude statistique portant sur un nombre important de langues. Ceci requiert de relier différentes sources d'informations sur les propriétés de l'environnement et sur les propriétés phonétiques des langues, et bénéficie grandement de l'emploi de systèmes d'information géographique (SIG).

Outre la fragilité des liens entre corrélations statistiques et causes réelles, une difficulté est la prise en compte des relations généalogiques et de contact entre langues. Ceci implique de recourir à des modèles complexes qui peinent encore à trouver leur place au sein de la linguistique. Des études distinctes, portant sur différents jeux de données, mettant en œuvre des approches diverses et complémentaires, sont donc les bienvenues pour dépasser les limites inhérentes à une étude isolée.

Les effets observés peuvent être faibles mais néanmoins significatifs. De nombreuses contraintes, en particulier sociales, s'exercent en effet sur les systèmes de sons et peuvent venir masquer l'adaptation acoustique à l'environnement.

## 2 Logique et hypothèses de l'étude

### 2.1 Végétation et adaptation acoustique

La communication humaine repose sur la transmission d'informations sonores entre un émetteur et un (ou plusieurs) récepteurs. Le canal de transmission peut modifier les propriétés d'un signal émis, et en particulier dégrader une partie de l'information qu'il contient. Ceci peut poser des problèmes de décodage de l'information par le récepteur, et entraîner des confusions.

Des études de perception de sons de la parole en milieu bruyant suggèrent que les voyelles résistent bien mieux au bruit que les consonnes [20]. Plus le bruit est important, plus il est difficile de distinguer les consonnes les unes des autres, avec des impacts différenciés selon leur position au sein de la syllabe.

Les populations humaines vivent dans des environnements qui se distinguent en particulier par leur végétation. La question se pose de l'impact de celle-ci sur la transmission de la parole, et d'une possible adaptation en retour, visible dans les inventaires de sons et de syllabes des langues du monde.

Nous faisons l'hypothèse qu'il existe une adaptation acoustique des langues humaines à la végétation, dans les traces de Maddieson [18] : une végétation plus dense, comme celle par exemple de la forêt équatoriale, dégrade

davantage la transmission des fréquences hautes des signaux de parole. Ces dernières composent la signature acoustique des consonnes, et en particulier celle des consonnes obstruantes. Dès lors, dans une végétation dense, pour contrer une mauvaise perception des consonnes, en particulier obstruantes, les changements linguistiques tendent à privilégier une diminution (ou une absence d'accroissement) du nombre de ces dernières, ainsi que des structures syllabiques complexes faisant intervenir un nombre important de consonnes. Maddieson a mis en évidence, pour plusieurs centaines de langues, un lien entre la latitude et un indice combinant la taille de l'inventaire consonantique et la complexité des syllabes pour une langue. Plus une langue est éloignée d'un environnement tempéré (vers 45° de latitude nord ou sud) - pour lequel on postule une végétation peu dense - moins son inventaire consonantique est riche et sa complexité syllabique importante.

On peut objecter que de nombreux locuteurs visent aujourd'hui dans des milieux urbanisés, où la végétation occupe peu de place. Toutefois, l'adaptation acoustique est vraisemblablement un phénomène s'exerçant sur le long terme, et qui a donc façonné les langues bien avant les récentes vagues de migrations des campagnes aux villes des populations du globe. Il est donc possible de laisser de côté en première approche cette question de l'urbanisation.

### 2.2 Tests statistiques géolinguistiques

Afin de tester notre hypothèse, nous disposons d'informations linguistiques et géographiques pour plusieurs centaines de langues du monde, réparties sur les différents continents et appartenant à différentes familles linguistiques. Il s'agit surtout des inventaires de consonnes et de voyelles pour ces langues, d'une mesure de la complexité moyenne de leurs syllabes, et de l'aire géographique de distribution des locuteurs. Nous disposons parallèlement de données écologiques de haute résolution spatiale pour l'ensemble des terres émergées, ce qui permet de calculer pour l'aire géographique d'une langue la valeur moyenne de la densité de végétation, de la température etc.

Notre but est d'établir un lien statistiquement significatif entre la densité moyenne de végétation dont les locuteurs d'une langue font l'expérience, et d'une part le nombre de consonnes différentes qu'ils emploient, d'autre part la complexité moyenne des syllabes avec lesquelles ils bâtissent leurs mots. Pour s'assurer de la pertinence d'un lien observé, il convient de prendre en compte d'autres facteurs pouvant expliquer en partie les variations de nos variables dépendantes (nombre de consonnes et complexité syllabique). Sur la base d'études précédentes, nous avons considéré le nombre de locuteurs, les relations historiques entre langues, les contacts linguistiques, la rugosité du terrain et la température [21,22,23].

## 3 Préparation des données

### 3.1 Source de données

En ce qui concerne les données linguistiques, nous nous sommes appuyés sur les données de la base "*Lyon-Albuquerque Phonological Systems Database*" (LAPSYD) [24], qui contient des informations sur la phonologie de plus de 750 langues. Certaines données sont manquantes, mais le soin apporté à l'homogénéisation des informations

fait de cette base une référence pour des études translinguistiques. Elle est en évolution permanente, et la version utilisée ici contient 763 langues.

Le “*World Language Mapping System*” (WLMS) [25] contient les aires géographiques sous forme vectorielle de plus de 7 000 langues, ainsi que leur affiliation et nombre de locuteurs. Nous avons approché pour chaque langue le degré de contact linguistique en calculant le recouvrement de son aire de distribution par d’autres langues. Nous avons également utilisé les aires de contact linguistique de la base *Autotyp* [26], qui correspondent à des régions du monde où les emprunts entre langues sont importants et bien établis.

L’établissement des liens entre les données du WLMS et les langues de LAPSUD s’est fait sur la base des codes iso et des noms des langues. 65 langues de LAPSUD ont été mises de côté car i) aujourd’hui disparues, ii) parlées sur des aires très importantes et écologiquement très variées, iii) sans correspondance dans le WLMS, principalement à cause de distinctions différentes entre dialectes et langues.

Pour les données environnementales, nous avons mis à profit des données issues de l’imagerie satellitaire. Nous avons eu recours à deux rasters pour l’altitude et le pourcentage de couverture arborée (pca par la suite) [27,28]. Ces deux jeux de données présentent la même très haute résolution spatiale – un point toutes les 15 secondes d’angle.



Figure 1: “Percent Tree Cover, Global Map v.2” [28]

Pour les données de température, nous avons enfin mis à profit un raster de valeurs du *WorldClim Global Climate Data*, d’une résolution spatiale de 30 secondes d’angle [29].

### 3.2 Opérations sur les données

Nous avons utilisé le logiciel libre QGIS pour le traitement des données géographiques. Nous avons tout d’abord dérivé un raster de rugosité du terrain à partir du raster d’altitude. Pour les aires de 698 langues, nous avons calculé les valeurs moyennes des variables environnementales. 10 langues ont été retirées du jeu de données car leur localisation géographique dans le WLMS n’était pas correctement couverte par les rasters de données. 2 autres langues, le français et le néerlandais, ont été éliminées à cause de leur distribution géographique sur plusieurs continents éloignés.

Sur les 686 langues restantes, 684 langues étaient renseignées pour le nombre de consonnes, 677 pour le nombre de consonnes sonantes et obstruantes, et 670 pour la complexité syllabique.

Une transformation logarithmique a été appliquée au nombre de locuteurs afin de se rapprocher d’une distribution normale.

Le logiciel R a été utilisé pour les analyses statistiques avec les packages *lme4* et *ordinal*, respectivement pour la réalisation de régressions de Poisson et de régressions ordinales.

## 4 Analyse des données

Une analyse de surface suggère des relations entre le pca et i) le nombre total de consonnes, ii) le nombre total de consonnes obstruantes, iii) la complexité syllabique. La figure 2 présente la première de ces trois relations et indique que le nombre de consonnes tend à diminuer avec l’accroissement du pourcentage de couverture arborée ( $R^2=12,2\%$ ). La figure 3 montre que les langues avec une complexité syllabique plus importante sont parlées dans les régions où le pca est plus faible.

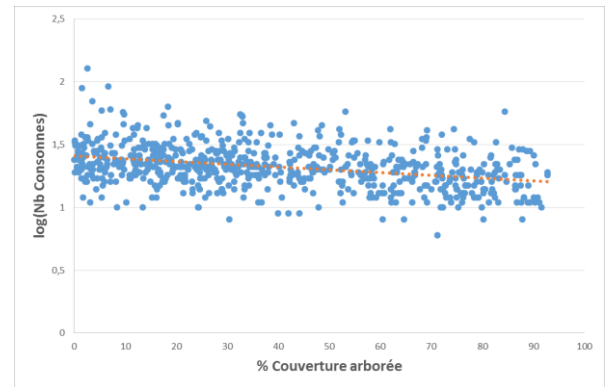


Figure 2: Relation entre le pca et le logarithme du nombre de consonnes pour 684 langues

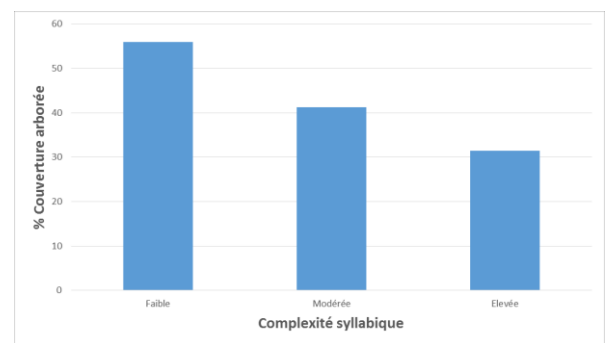


Figure 3: Relation entre la complexité syllabique et le pourcentage moyen de couverture arborée pour chaque catégorie de complexité syllabique (670 langues au total)

Pour prendre en compte d’autres facteurs pouvant rendre compte partiellement du nombre de consonnes ou de la complexité syllabique, nous avons eu recours à des modèles de régression multiple. Pour chacun de ces modèles, le pca, la rugosité du terrain, la température, le degré de contact linguistique et le nombre de locuteurs ont joué le rôle d’effets fixes. Ces variables ont été réduites et centrées pour faciliter la convergence des modèles. Deux effets aléatoires ont été également inclus à chaque fois : la généalogie linguistique (avec un effet de sous-famille imbriqué dans un effet famille), et l’appartenance à une aire de contact linguistique. L’affiliation à une famille linguistique ou à une zone de contact peut expliquer la présence d’un faible ou d’un fort nombre de phonèmes par héritage ou emprunt ; il convient donc de s’assurer que la végétation joue un rôle même quand ces phénomènes sont pris en compte, de même que les effets d’autres variables indépendantes (rugosité, température etc.)

Différents modèles ont été construits et analysés :

- Régression de Poisson (RP) du nombre de consonnes : effet très significatif du pca ( $\beta_{pca} < 0$ ) et du degré de contact ( $\beta_{dc} > 0$ ), effet significatif de la température ( $\beta_T < 0$ )
- RP du nombre de voyelles : aucun effet significatif
- RP du nombre de consonnes obstruantes : effet très significatif du pca ( $\beta_{pca} < 0$ ), effet significatif du degré de contact ( $\beta_{dc} > 0$ ) et de la température ( $\beta_T < 0$ )
- RP du nombre de consonnes sonantes : effet significatif du pca ( $\beta_{pca} < 0$ )
- Régression ordinale de la complexité syllabique (avec 3 niveaux : simple, modérée, élevée) : effet significatif du pca et de la température ( $\beta_{pca} < 0$  et  $\beta_T < 0$ )

Effets fixes	$\beta$	Erreur standard	z value	Pr(> z )
Intercept	1.88	.037	51.18	< 2e-16
Tree	-.050	.021	-2.34	<b>.019</b>
Rugosity	.014	.021	.67	.50
T°	-6e-4	.025	-.02	.98
Degree of contact	-.010	.019	-.54	.59
Log(Nb of speakers)	.002	.013	.12	.90

Tableau 1: Coefficients de la régression de Poisson du nombre de consonnes

Effets fixes	$\beta$	Erreur standard	z value	Pr(> z )
Intercept	3.04	.039	78.10	< 2e-16
Tree	-.056	.014	-4.04	<b>5.39e-05</b>
Rugosity	.018	.014	1.30	.19
T°	-.050	.020	-2.45	<b>.014</b>
Degree of contact	.034	.009	3.71	<b>2.03e-04</b>
Log(Nb of speakers)	.0006	.013	.04	.97

Tableau 2: Coefficients de la régression de Poisson du nombre de consonnes obstruantes

Effets fixes	$\beta$	Erreur standard	z value	Pr(> z )
Intercept	2.65	.047	56.13	< 2e-16
Tree	-.059	.017	-3.46	<b>5.42e-04</b>
Rugosity	.022	.017	1.31	.19
T°	-.065	.025	-2.59	<b>9.56e-03</b>
Degree of contact	.034	.011	3.28	<b>1.05e-03</b>
Log(Nb of speakers)	-.006	.017	-.36	.72

Tableau 3: Coefficients de la régression de Poisson du nombre de consonnes sonantes

Il apparaît ainsi qu'un effet significatif de la végétation s'exerce sur les consonnes mais pas les voyelles, et parmi les consonnes légèrement sur les sonantes et de façon plus marquée sur les obstruantes. Ces résultats se complètent pour soutenir l'idée que les sons basés sur de hautes fréquences sont défavorisés quand la végétation est dense.

On observe aussi qu'une température plus élevée semble pénaliser les consonnes obstruantes mais pas les consonnes sonantes (et pénalise l'ensemble des consonnes en conséquences, à l'inverse des voyelles). Une hypothèse intéressante qui demanderait des investigations plus poussées est qu'un air plus chaud dégrade aussi la transmission des fréquences élevées du signal de parole à cause de la présence de colonnes de turbulences.

## 5 Conclusion

Les différents modèles employés suggèrent clairement un effet du pourcentage de couverture arborée sur le nombre de consonnes et en particulier de consonnes obstruantes dans les inventaires phonologiques, ainsi que sur la complexité syllabique. Même si la taille de cet effet n'est pas très importante, sa significativité est un argument en faveur d'une adaptation acoustique de la communication humaine à l'environnement.

Différentes prolongations de cette étude sont envisagées, comme sa réplication avec d'autres bases de données linguistiques disponibles, ou une sélection plus poussée des langues étudiées, selon un critère d'homogénéité écologique de leur aire de distribution.

## Remerciements

Les auteurs remercient le Labex ASLAN (ANR-10-LABX-0081) de l'Université de Lyon pour son soutien financier dans le cadre du programme "Investissements d'Avenir" (ANR-11-IDEX-0007) du gouvernement français, opéré par l'Agence Nationale de la Recherche.

## Références

- [1] C. M. Harris, Absorption of Sound in Air versus Humidity and Temperature, *JASA* **40**, 148–59 (1966).
- [2] J. E. Piercy, T. F. W. Embleton, and L. C. Sutherland, Review of noise propagation in the atmosphere, *JASA* **61**(6), 1403–1418 (1977).
- [3] D. G. Richards and R. H. Wiley, Reverberations and Amplitude Fluctuations in the Propagation of Sound in a Forest: Implications for Animal Communication, *American Naturalist* **115**(3), 381–399 (1980).
- [4] F. Yang, Z. Bao, Z. Zhu, and J. Liu, The Investigation of Noise Attenuation by Plants and the Corresponding Noise Reducing Spectrum, *J. of Env. Health* **72**(8), 8–15 (2010).
- [5] Peters, G. and M. K. Peters, Long-distance call evolution in the Felidae: effects of body weight, habitat, and phylogeny, *Biological J. of the Linnean Society* **101**, 487–500 (2010).
- [6] R. H. Wiley and D. G. Richards, Physical constraints on acoustic communication in the atmosphere: implications for the evolution of animal vocalizations, *Behav. Ecol. Sociobiol.* **3**, 69–94 (1978).
- [7] E. S. Morton, Ecological Sources of Selection on Avian Sounds, *American Naturalist* **109**(965), 17–34 (1975).
- [8] E. Ey and J. Fischer, The “Acoustic Adaptation Hypothesis” - A Review of the Evidence from Birds, Anurans and Mammals, *Bioacoustics* **19**(1-2), 21–48 (2009).
- [9] C. Schneider, J. K. Hodges, J. Fischer, and K. Hammerschmidt, Acoustic Niches of Siberut Primates, *Int. J. of Primatology* **29**(3), 601–613 (2008).
- [10] N. Mathevon, T. Aubin, J. Vielliard, M. L. da Silva, F. Sebe, and D. Boscolo, Singing in the rain forest: How a tropical bird song transfers information. *PLoS ONE* **3**(2), 2–7 (2008).
- [11] H. Slabbekoorn and M. Peet, Ecology: Birds sing at a higher pitch in urban noise, *Nature* **424**(6946), 267–267 (2003).
- [12] W. Wood and S. Yezerinac, Song sparrow (*melospiza melodia*) song varies with urban noise, *The Auk* **123**, 650–659 (2006).
- [13] D. Nettle, Language Diversity in West Africa: An Ecological Approach, *J. of Anthropological Archaeology* **15**, 403–438 (1996).
- [14] R. L. Munroe, R. H. Munroe, and S. Winters, Cross-cultural correlates of the consonant-vowel (CV) syllable, *Cross-Cultural Research* **30**, 60–83 (1996).
- [15] C. Coupé, Commentary: Defining and assessing constraints on linguistic forms, *J. of Language Evolution* **1**(1), 52–55 (2016).
- [16] M. Fortescue, *Orientation Systems of the North Pacific Rim*, Museum Tusulanum Press, Copenhagen (2011).
- [17] B. Palmer, Pointing at the lagoon: Directional terms in Oceanic atoll-based languages, in *Language Description, history and development. Linguistic indulgence in memory of Terry Crowley*, J. Siegel, J. Lynch, and D. Eades (Eds), John Benjamins, Amsterdam / Philadelphia, 101–118 (2007).
- [18] C. Everett, Evidence for Direct Geographic Influences on Linguistic Sounds: The Case of Ejectives, *PLoS ONE* **8**(6), e65275 (2013).
- [19] C. Everett, D. E. Blasi, and S. G. Roberts, Climate, vocal folds, and tonal languages: Connecting the physiological and geographic dots. *PNAS* **112**(5), 1322–1327 (2015).
- [20] J. Meyer, L., Dentel, and F. Meunier, Speech recognition in natural background noise, *PLoS ONE* **8**(11) (2013).
- [21] I. Maddieson, Phonological complexity in linguistic patterning, in *Proc. of the 17th ICPHS*, 17–21 (2011).
- [22] J. Hay and L. Bauer, Phoneme inventory size and population size, *Language* **83**, 388–400 (2007).
- [23] S. Moran, D. McCloy, and R. Wright, Revisiting population size vs. phoneme inventory size, *Language* **88**(4), 877–893 (2012).
- [24] I. Maddieson, S. Flavier, E. Marsico, C. Coupé, and F. Pellegrino, LAPSyD : Lyon-Albuquerque Phonological Systems Database, in *Proc. of the 14th Interspeech Conference* (2103).
- [25] Global Mapping International and SIL International, *World Language Mapping System, Language area and point data for Geographic Information Systems (GIS)*, [www.worldgeodatasets.com/language](http://www.worldgeodatasets.com/language) (2012).
- [26] B. Bickel and J. Nichols, *The Autotyp Project*, <http://www.autotyp.uzh.ch> (2001).
- [27] Geospatial Information Authority of Japan, *Elevation, Global Map V.2*, [www.iscgm.org/gm/el.html](http://www.iscgm.org/gm/el.html) (2013).
- [28] Geospatial Information Authority of Japan, Chiba University, and collaborating organizations, *Vegetation (Percent Tree Cover), Global Map V.2*, [www.iscgm.org/gm/ptc.html](http://www.iscgm.org/gm/ptc.html) (2013).
- [29] R. J. Hijmans, S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis, Very high resolution interpolated climate surfaces for global land areas, *Int. J. Climatology* **25**, 1965–1978 (2005).