



GesturalOrigins : Un cadre ascendant pour l'établissement de données gestuelles systématiques sur les espèces de singes

Charlotte Grund¹ - Gal Badihi¹ - Kirsty E. Graham¹ - Alexandra Safryghin¹ - Catherine Hobaiter¹

Accepté : 3 février 2023 / Publié en ligne : 15 mars 2023
© Le(s) auteur(s) 2023

Résumé

Les méthodologies actuelles présentent des obstacles importants à la compréhension des modèles de communication gestuelle des individus, des populations et des espèces. Pour résoudre ce problème, nous présentons un cadre de collecte de données ascendantes pour l'étude des gestes : GesturalOrigins. Par "ascendant", nous entendons que nous minimisons les choix structurels *a priori*, ce qui permet aux chercheurs de définir des concepts plus larges (tels que les "types de gestes", les "latences de réponse" ou les "séquences de gestes") de manière flexible une fois le codage terminé. Les données peuvent être facilement réorganisées pour permettre la reproduction et la comparaison avec un large éventail d'ensembles de données dans les analyses publiées et planifiées. Nous présentons des paquets, des modèles et des instructions pour l'ensemble du processus de collecte et de codage des données. Nous illustrons la flexibilité offerte par notre outil méthodologique à l'aide d'exemples concrets de communication gestuelle (de grands singes), en démontrant les différences de durée des phases d'action entre les différents types d'action gestuelle et en montrant comment les variations entre les espèces dans la latence de réponse aux demandes gestuelles peuvent être révélées ou masquées par les choix méthodologiques. Bien que GesturalOrigins soit construit à partir d'une perspective centrée sur le singe, le cadre de base peut être adapté à une gamme d'espèces et potentiellement à d'autres systèmes de communication. En rendant nos méthodes de codage gestuel transparentes et en libre accès, nous espérons permettre une comparaison plus directe des résultats entre les groupes de recherche, améliorer les collaborations et faire progresser le domaine en abordant certaines questions de longue date dans la recherche comparative sur les gestes.

Mots-clés Codage vidéo - Phases d'action du geste - Origines gestuelles - Communication visuelle - Évolution du langage

Introduction : Établir des données systématiques dans la recherche comparative sur les gestes des grands singes

Les signaux gestuels sont utilisés dans un large éventail d'espèces (par exemple, les babouins : Molesti et al., 2020 ; mangabeys : Schel et al., 2022 ; macaques : Gupta & Sinha, 2019 ; corbeaux : Pika & Bugnyar, 2011 ; méroü : Vail et al., 2013), mais la plupart des recherches sur les gestes se sont concentrées sur les grands singes (par exemple, les chimpanzés : Tomasello et al., 1985 ; Liebal et al., 2006 ; Hobaiter & Byrne, 2011a ; bonobos : Pika et al., 2005 ; Graham et al., 2017 ; gorilles : Pika et al., 2003 ; Genty et al., 2009 ; orangs-outans : Liebal et al., 2006 ; Cartmill & Byrne, 2007). Bien que travaillant sur les mêmes espèces ou sur des espèces étroitement apparentées, une gamme variée de méthodes d'étude des gestes des grands singes a vu le jour au fil du temps.

les quatre dernières décennies. La variation dans l'approche, y compris les différentes méthodes, les cadres de recherche et les espèces, est essentielle pour l'exploration robuste du comportement (Rodrigues et al., 2021), mais la diversité des approches pour étudier les gestes des grands singes a soulevé des obstacles pour la comparabilité entre les études, et des examens récents ont souligné la nécessité d'une plus grande transparence méthodologique (par exemple, Scott & Pika, 2012 ; Fröhlich & Hobaiter, 2018 ; Bourjade et al., 2020 ; Rodrigues et al., 2021). Notre capacité à aborder les questions comparatives ne dépend pas de l'adoption d'une approche universelle, mais elle dépend de méthodes suffisamment transparentes et détaillées pour permettre la comparaison directe de données similaires. Compte tenu de l'ampleur de la question et de la nature de l'étude d'individus à longue durée de vie, au comportement flexible et hautement social, le moyen le plus simple et le plus efficace d'étudier le comportement gestuel chez les grands singes et les espèces similaires est de collaborer et de comparer les résultats des différentes études.

Ce défi n'est pas propre à la recherche sur les gestes des singes : des travaux récents sur la "crise de la répliation" (Open Science Collaboration, 2015 ; Baker, 2016) mettent en évidence les problèmes de fiabilité des résultats dans la recherche comportementale (Webster & Rutz, 2020 ; Rutz & Webster, 2021). Bien qu'il puisse parfois

✉ Charlotte Grund grund.charlotte@gmx.de

Catherine Hobaiter
clh42@st-andrews.ac.uk

¹ École de psychologie et de neurosciences, Université de St Andrews, Fife, Écosse KY16 9JP, Royaume-Uni

Il est impossible ou inapproprié de reproduire exactement les données (Farrar et al., 2021), mais la variabilité devient un problème lorsque les sources de variation sont opaques - sans méthodes transparentes, nous sommes incapables de discerner si les différences dans les résultats proviennent, par exemple, de biais d'échantillonnage ou d'erreurs méthodologiques (Schweinfurth & Call, 2021). L'un des moyens de relever ce défi consiste à améliorer la transparence dans la description des données - par exemple, le cadre STRANGE clarifie la nature des biais d'échantillonnage dans la recherche animale (Webster & Rutz, 2020) en adoptant une approche de type "conception, déclaration et discussion". En tandem, les cadres méthodologiques ascendants permettent aux chercheurs de construire et de reconstruire des ensembles de données de différentes manières, offrant à la fois une flexibilité individuelle et la possibilité d'une comparaison similaire (lorsque les données brutes sont disponibles ; Eaton et al., 2018). Nous visons à contribuer à une augmentation de la robustesse méthodologique et de la transparence dans la recherche sur les gestes en présentant un cadre de collecte de données ascendant : GesturalOrigins. Nous visons à fournir suffisamment de détails pour que ce cadre puisse être facilement adopté par d'autres chercheurs. Nous intégrons et nous appuyons sur des décennies de progrès méthodologiques dans la recherche sur les gestes et nous présentons des paquets, des modèles et des instructions pour l'ensemble du processus de collecte de données : de la collecte de données vidéo interactionnelles au codage des événements de communication gestuelle dans ELAN, un logiciel libre accessible (ELAN, 2022), en passant par l'exportation des données codées et leur transformation en vue de leur stockage dans une base de données appropriée, par exemple Filemaker. Nous utilisons des exemples pratiques de communication gestuelle chez le chimpanzé et le gorille pour démontrer l'adaptabilité de notre cadre de travail, en fournissant un tutoriel vidéo et des protocoles pour la procédure de conservation des données (collecte, codage et exportation) qui décrivent également le processus de prise de décision théorique impliqué. En tant que primatologues, nous adoptons une perspective centrée sur les primates, mais le cadre de base peut potentiellement être adapté à toute une série d'espèces (par exemple, les éléphants de la savane africaine, voir la page d'accueil du laboratoire Wild Minds <https://www.wildminds.ac.uk/> pour plus d'informations). projets en cours).

Le cadre de codage GesturalOrigins

Considérations de base pour la mesure du comportement gestuel

Les études d'observation du comportement animal fournissent des descriptions riches et nuancées de la manière dont les individus, les groupes et les espèces interagissent avec leur environnement physique et social (Altmann, 1974). Cependant, l'observation directe en personne est vulnérable à une série de biais d'attention et de perception (Kaufman & Rosenthal, 2009). Les données gestuelles comprennent des variables qui sont particulièrement sujettes aux erreurs d'observation directe, car elles reposent sur des durées de l'ordre de la seconde (par exemple, dans le cas de l'attente d'une réponse, tableau 1, section suivante) ; sur des différences subtiles dans Le comportement gestuel peut être mesuré à l'aide d'une série d'indicateurs de comportement

(par exemple, les changements dans la direction du regard ("attention visuelle"), tableau 1, section suivante) ou de mouvements d'action (par exemple, pour déterminer la différence entre un "swing" et un "fling" dans une interaction sociale où beaucoup de choses se passent en même temps) ; il implique souvent la collecte de points de données provenant de plusieurs individus qui peuvent se chevaucher dans le temps. En particulier dans des conditions naturalistes où l'observation est encore plus gênée par des environnements visuellement denses et un éclairage variable, le comportement gestuel ne peut pas être recueilli de manière fiable au cours de l'échantillonnage focal en direct. Le codage basé sur la vidéo, considéré comme un étalon-or dans la recherche observationnelle (Gilmore & Adolph, 2017), fournit non seulement des mesures temporelles précises et l'identification d'actions subtiles et évidentes, mais permet également aux chercheurs de revoir les interactions, d'améliorer la formation des nouveaux codeurs et de faciliter la réalisation précise d'importants tests de fiabilité intra-codeur et inter-codeur (Burghardt et al. 2012). L'annotation dans la vidéo, où les données codées sont directement liées au matériel vidéo, augmente encore la précision et allège le fardeau que représente le fait de revisiter les données pour évaluer la fiabilité ou générer des comparaisons. Il existe plusieurs programmes de codage comportemental en vidéo, par exemple BORIS (Friard & Gamba, 2016), Observer XT (Noldus et al., 2000), Solomon Coder (Péter, 2011) ou ELAN (ELAN, 2022), pour n'en citer que quelques-uns. Inspirés par des recherches antérieures sur la communication basée sur la vidéo (par exemple, Heesen et al., 2020), nous avons mis en œuvre notre système de codage dans le logiciel libre d'annotation vidéo linguistique ELAN (ELAN, 2022). Nous décrivons étape par étape comment coder avec le schéma de codage GesturalOrigins dans notre protocole de codage (GOv1.0_Protocol), accessible par le biais du matériel électronique supplémentaire (ESM) de ce manuel script sur GitHub (https://github.com/CharlotteGrund/GesturalOrigins_Coding-methods_paper, ainsi que tous les autres fichiers référencés comme ESM tout au long du manuscrit). Outre des descriptions détaillées de toutes nos variables, le protocole comprend des sections sur la manière d'exporter les données codées d'ELAN et propose des paquets R qui convertissent le fichier d'exportation en un fichier de données "propre" pour une analyse plus approfondie.

Qu'est-ce qu'un comportement gestuel ? Indicateurs comportementaux de l'intention

Conformément à un grand nombre de recherches sur les gestes des animaux non humains (Graham et al., 2017 ; Gupta & Sinha, 2019 ; Hobaiter & Byrne, 2011a ; Liebal et al., 2006 ; Molesti et al., 2020 ; Pika & Bugnyar, 2011 ; Schel et al., 2022 ; Tomasello et al., 1985 ; Vail et al., 2013), nous définissons un geste comme un signal produit avec le corps, qui est un "mouvement physique mécaniquement inefficace" (Hobaiter & Byrne, 2011a) d'une ou plusieurs parties du corps ou du corps dans son ensemble, utilisé intentionnellement pour atteindre un objectif social spécifique (Bard, 1992 ; Leavens & Hopkins 1998 ; Pika et al., 2003 ; Pollick & de Waal, 2007 ; Tomasello et al., 1985 ; Hobaiter & Byrne, 2011a). Un geste est produit volontairement par l'individu qui le signale (c'est-à-dire qu'il ne s'agit pas d'une réaction réflexe ou automatique à un indice externe ou interne) et est

visant à susciter une réponse comportementale particulière de la part d'un individu spécifique (le destinataire). La communication est réussie lorsque le destinataire modifie son comportement d'une manière qui représente un objectif plausible du point de vue de l'émetteur du signal (le résultat apparemment satisfaisant, généralement une forme d'interaction sociale) et que l'émetteur du signal cesse de signaler (indiquant qu'il est apparemment satisfait du changement de comportement). Un geste peut impliquer une manipulation mécanique du destinataire par le signaleur (par exemple, dans un geste de "poussée" - voir ESM : GOv1.0_Gesture_action_definitions.xlsx pour une définition des actions gestuelles), mais jamais dans une mesure qui permette d'atteindre le but lui-même (l'action est *mécaniquement inefficace* : une "poussée" peut être utilisée pour signaler au destinataire de déplacer son corps dans une certaine direction, mais la force utilisée ne doit pas être efficace pour déplacer le corps du destinataire vers l'endroit désiré).

Les chercheurs sur les gestes des animaux non humains ont adapté les critères d'intentionnalité pour la communication non humaine à partir de la recherche sur l'intentionnalité dans la communication préverbale des enfants humains et leur passage d'actes de communication perlocutoire à illocutoire (Bates et al. 1975, résumé dans le tableau 1 ci-dessous). Bien que la présence de marqueurs individuels tels que ceux-ci puisse être expliquée d'autres manières, leur utilisation combinée et régulière permet d'affirmer que la communication des signaleurs est mieux comprise comme intentionnelle.

Dans notre routine de codage, nous présélectionnons tous les clips vidéo collectés pour un comportement gestuel potentiel (généralement lorsque deux individus (ou plus) se trouvent à moins de 5 m l'un de l'autre et montrent une certaine forme d'engagement social) et appliquons les critères d'intentionnalité décrits dans le tableau 1 pour trier les exemples réels d'utilisation de gestes en vue d'un codage détaillé. Bien que nous nous concentrons ici sur les signaux gestuels clairement intentionnels dans les interactions dyadiques, le système de codage GesturalOrigins peut également être appliqué à des exemples d'utilisation de gestes qui semblent cibler plusieurs individus ou lorsque les critères d'utilisation intentionnelle ne sont pas clairs. Cette ouverture signifie qu'il n'est pas nécessaire de décider *a priori* de ne coder que les exemples intentionnels dirigés vers un but - ces derniers peuvent au contraire être sous-évalués à partir des données par la suite.

en fonction des critères qui sont ensuite appliqués. Ce type de codage augmente considérablement l'effort de codage nécessaire pour détecter les gestes, mais il peut s'avérer intéressant dans le cadre d'études visant à déterminer comment les actions gestuelles intentionnelles varient par rapport à d'autres signaux physiques ou à étudier les différents degrés d'intentionnalité dans la production de signaux (par exemple, au cours de l'ontogenèse). De plus, cela permet également d'explorer l'utilité de ces marqueurs pour établir l'utilisation intentionnelle à travers les types de gestes (par exemple, il existe un biais établi vers les signaux visuels (Dafreville et al., 2021)). L'amélioration des marqueurs d'intentionnalité pour les différentes modalités dans lesquelles les gestes et d'autres types de signaux se produisent est un défi scientifique (par exemple, Townsend et al., 2017 ; Fröhlich & Hobaiter, 2018), et nous espérons que les schémas de codage ascendants s'avéreront des outils précieux pour les approches axées sur les données afin de relever certains de ces défis.

Quelle unité constitue un geste ? Décrire les instances du geste (forme physique)

S'il existe un large consensus sur la manière de définir le comportement gestuel et d'appliquer ses marqueurs comportementaux, les méthodologies utilisées pour identifier et décrire les formes physiques des signaux gestuels varient davantage (Rodrigues et al., 2021 ; Hobaiter & Byrne, 2017). Les études varient quant aux actions ou aux parties du corps qu'elles prennent en compte ; certaines incluent les postures du corps entier (par exemple, Hobaiter & Byrne, 2011a ; Molesti et al., 2020) et/ou les "gestes faciaux" (par exemple, Cartmill & Byrne, 2010 ; Molesti et al., 2020), tandis que d'autres se limitent uniquement aux mouvements des mains et des bras (par exemple, Pollick & de Waal, 2007 ; Roberts et al., 2012). Dans une récente revue systématique de la littérature sur les gestes humains et non humains, Rodrigues et al. (2021) ont constaté que dans 163 études, 54% des chercheurs n'incluaient que des mouvements manuels, 34% incluaient à la fois des mouvements manuels et non manuels, et le reste ne spécifiait pas la partie du corps ou n'incluait que des gestes non manuels. Les approches de la construction de types de gestes diffèrent également dans la façon dont les

Tableau 1 Définitions des indicateurs comportementaux utilisés pour identifier la signalisation

gestuelle intentionnelle	Terme	Définition
Vérification de l'attention		Le signaleur vérifie l'état de l'attention visuelle du destinataire avant de produire le signal et adapte sa signalisation en conséquence (par exemple, en utilisant des signaux uniquement visuels lorsque le destinataire regarde et des signaux sonores ou tactiles lorsqu'il ne regarde pas, en augmentant les chances qu'un signal soit perçu et en faisant preuve d'une certaine "sensibilité à l'état de l'attention").
Attente de la réponse		Le signaleur fait une pause et attend que le destinataire réponde à sa demande (l'indice comportemental est ici une pause dans les gestes et la surveillance visuelle du destinataire).
Persistance du but		Le signaleur continue de signaler lorsque le destinataire ne répond pas et persiste et/ou développe ses gestes jusqu'à ce que le but soit atteint.
Inefficacité mécanique	le geste du signaleur	peut manipuler (mécaniquement) le destinataire, mais jamais au point d'atteindre le but lui-même (l'action gestuelle est <i>mécaniquement inefficace</i> : une "poussée" peut être utilisée pour signaler au destinataire de déplacer son corps dans une certaine direction, mais la force utilisée ne devrait pas être efficace pour déplacer le corps du destinataire à l'endroit désiré).

Certaines études classent les actions produites avec une ou deux mains comme différents "types de gestes", par exemple Genty et al., (2009) ou dans la mesure où elles discriminent les aspects fonctionnels d'un geste qui sont indépendants de son apparence morphologique (par exemple, "toiletage présent" et "sexuel présent" peuvent impliquer exactement le même mouvement dans différents contextes comportementaux ; par exemple, Hobaiter & Byrne, 2011a). Ainsi, malgré les éléments communs généralement utilisés pour décrire un "type" de geste : actions, parties du corps, implication d'objets, utilisation de la répétition (Hobaiter & Byrne, 2017), les chercheurs les recombinaient souvent de manière non systématique d'une étude à l'autre.

Ensemble, ces incohérences méthodologiques peuvent aboutir à des résultats divergents entre les études en ce qui concerne les types de gestes, la taille des répertoires et le nombre de gestes observés, même au sein d'un ensemble identique de vidéos comportementales de grands singes. Dans certains cas, il est encore possible de faire des comparaisons en regroupant (par exemple, en combinant toutes les actions unimanuelles et bimanuelles), mais ce faisant, on risque de perdre une résolution importante. Par exemple, l'exclusion des gestes effectués avec le corps ou des parties du corps autres que les membres réduirait le répertoire des chimpanzés décrit par Hobaiter et al. (2011a) d'environ 26 %. Il est important de noter qu'il n'y a pas de "bonne" façon d'analyser les gestes, mais lorsque chaque schéma de codage est adapté aux questions particulières d'intérêt, les données codées brutes sont biaisées d'une manière qui rend les comparaisons entre les études difficiles. Le potentiel de comparaison entre les études est particulièrement limité lorsque les définitions ne sont pas transparentes ou que la construction des types de gestes ne suit pas des règles cohérentes. Notre approche relève ce défi en codant les unités gestuelles de bas en haut, ce qui permet aux chercheurs de décrire les "types" de gestes à différents niveaux et à l'aide d'approches probabilistes fondées sur des données après le codage.

La base de la description des unités gestuelles dans notre codage est la suivante : "La description des unités gestuelles est la base de la description des unités gestuelles dans notre codage.

Le premier niveau est l'*action gestuelle*, le mouvement corporel qui décrit l'instance gestuelle en cours (par exemple, un "reach" est une extension coordonnée vers un destinataire, tandis qu'un "beckon" est un mouvement d'écopage qui s'éloigne du signaleur et le ramène vers lui ; voir ESM : GOv1.0_Gesture_action_definitions.xlsx pour une liste complète des actions gestuelles et de leurs définitions). Nous codons ensuite plusieurs variables en tant que niveaux "enfants" de l'action gestuelle principale (c'est-à-dire des variables liées à l'annotation de l'enregistrement du geste, le niveau "parent") qui décrivent plus en détail la production physique de l'instance gestuelle focale (voir la figure 1 ci-dessous pour les principales variables codées pour chaque instance d'action gestuelle) :

Ainsi, chaque instance de geste comporte le mouvement d'action en son centre ainsi que des caractéristiques supplémentaires ("modificateurs"), qui peuvent être utilisées comme éléments de construction pour spécifier davantage le comportement observé, ce qui permet une certaine flexibilité dans la construction des "types" de gestes (et des répertoires qui en résultent) à différents niveaux de résolution après le codage. Nous décrivons les formes gestuelles résultantes, la combinaison non aléatoire d'actions gestuelles et de leurs modificateurs (voir figure 1), comme des "morphes", sans supposer a priori qu'elles reflètent ou non un comportement particulier.

type" ou autre catégorie de geste. La création de "morphes" gestuels peut se faire soit manuellement (comme cela a été fait précédemment, mais sans la flexibilité et la transparence de l'approche actuelle), soit par l'application de nouvelles méthodes informatiques aux données gestuelles codées (par exemple, algorithmes de regroupement ; Mielke et al. *en prép.*). Des méthodes d'apprentissage automatique similaires ont été utilisées pour décrire les répertoires vocaux (par exemple, Keen et al., 2021), détectant des modèles structurels d'ordre supérieur potentiellement significatifs dans le comportement communicatif d'une espèce que nous, en tant qu'observateurs humains (ne pouvant examiner que chaque geste à la fois), aurions pu manquer (Wadewitz et al., 2015). Bien que nous pensions que de nombreuses actions gestuelles que nous décrivons (ici et dans notre matériel complémentaire GOv1.0_Gesture_action_definitions.xlsx) sont transférables à des espèces autres que le singe, et en particulier à d'autres espèces de primates, certaines d'entre elles sont typiques de l'espèce. Les chercheurs ont la possibilité d'ajuster facilement la liste des actions gestuelles pour tenir compte de tout mouvement gestuel typique de l'espèce (comme c'est le cas pour tout autre "vocabulaire contrôlé" présenté ; par exemple, notre groupe de recherche a récemment adapté ce schéma pour l'utiliser chez les éléphants de la savane africaine - certaines des adaptations comprenaient l'ajout des oreilles et de la trompe à la liste des parties du corps).

Durée du geste et unités d'action informatives

La description des types de gestes (et des répertoires qu'ils constituent) n'est qu'une première étape, bien qu'essentielle, dans la description du fonctionnement et de l'action des gestes dans les interactions communicatives. Nous considérons chaque geste dans un événement de communication comme un élément porteur d'une certaine forme d'"information" que le signaleur souhaite transmettre au destinataire (potentiellement modifiée et/ou spécifiée en conjonction avec d'autres facteurs tels que le contexte dans lequel il est produit, la relation signaleur/destinataire ou sa combinaison avec d'autres actes de communication ; Graham et al., 2022). Nous entendons par là que l'action gestuelle du signaleur contient des informations sur l'objectif visé, une raison pour laquelle le signaleur a choisi cette action gestuelle particulière pour cet objectif de communication particulier dans cette situation particulière, et qu'il existe diverses autres sources d'"informations" inhérentes à la situation (par exemple, le contexte ou le rang du signaleur) qui peuvent modifier/compléter les informations véhiculées par l'action gestuelle elle-même. Mais quelle est l'unité de l'instance gestuelle elle-même qui porte l'information?

Les études sur les gestes de la co-parole humaine distinguent quatre étapes d'un geste : la préparation, le coup d'action, le maintien et la récupération (Kendon, 2004). Il existe des distinctions potentiellement importantes entre les informations contenues dans le mouvement d'action lui-même (phase d'action) et la poursuite de l'action (phase de maintien ou de répétition). Dans une communication, la phase d'action (par exemple, tendre le bras vers le destinataire dans un geste d'extension) peut être suffisante pour que le destinataire comprenne le geste, tandis que la phase (facultative) de maintien/ répétition (c'est-à-dire le maintien de l'action du geste d'extension) peut ajouter d'autres informations à la communication.

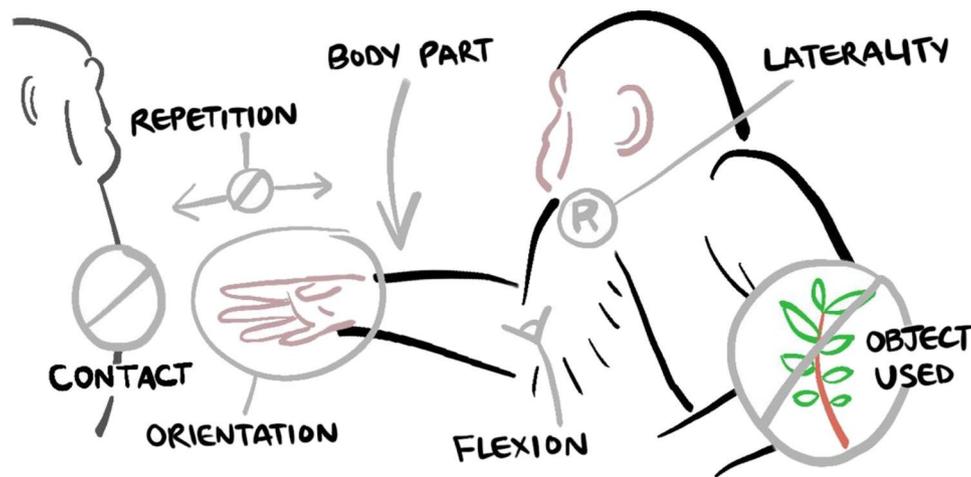


Fig. 1 Enregistrement du geste : codage de l'action du geste et de ses modificateurs - exemple : "atteindre". *Remarque.* Construction ascendante du type de geste : codage de l'action gestuelle et des modificateurs qui décrivent la production physique de l'action gestuelle de manière plus détaillée, en prenant l'exemple illustré de "reach". **Enregistrement du geste** (annotation parentale) : quel mouvement corporel est effectué (*ici* : "atteindre") ? **Signaleur de la partie du corps** : quelle partie du corps a été utilisée (*ici* : "Bras") ? **Latéralité du signaleur** : s'agit-il de la partie gauche ou droite du corps (*ici* : "Droite") ? **Objet utilisé** : le geste a-t-il été produit à l'aide d'un objet (*ici* : "Aucun") ? **Flexion** : dans le cas d'un geste effectué par un membre libre (sans contact avec un objet, un substrat ou une autre partie du corps), nous examinons si le coude, le poignet et/ou les doigts ont été pliés à plus de 45° (*ici* : "Coude") ? **Orientation** : dans le cas d'un geste d'un membre libre, dans quelle direction la paume de la main ou la plante du pied est-elle orientée (*ici* : "Côté") ? **Nombre de répétitions** : pour les types de gestes qui ont un rythme

mouvements répétés, combien de fois l'action a-t-elle été répétée (*ici* : Pas de valeur car le "reach" n'est pas une action gestuelle répétée. Notez que ce niveau est une variable de *texte libre*) ? **Contact avec la partie du corps** : pour les gestes de contact, nous codons également la partie du corps du destinataire avec laquelle le geste entre en contact (*ici* : "Aucune" car "atteindre" n'est pas une action de geste de contact). En outre, nous estimons l'*audibilité* de l'instance du geste et s'il existe une preuve de *directionnalité*, par exemple, si la main a été tendue vers un individu ou un lieu d'intérêt potentiel (Note : ces deux variables ne sont pas illustrées ici). Voir le fichier Excel du vocabulaire contrôlé (GOv1.0_Elan_controlled_vocabulary.xlsx) pour obtenir la liste complète des options pour les variables d'action et de modification du geste et le protocole de codage de GesturalOrigins (GOv1.0_Protocol) pour obtenir des détails sur la manière de coder chaque variable (les deux fichiers sont accessibles *ici* : https://github.com/CharlotteG_rund/Gestural_Origins_Coding_methods_paper).

des informations, par exemple, sur la volonté du signaleur d'attendre ou de persister. En gardant ce cadre à l'esprit, nous distinguons deux grandes catégories de types d'actions gestuelles : celles qui intègrent une phase facultative de maintien/répétition (par exemple, comme dans les actions gestuelles "reach" ou "raise" (phase facultative de maintien) ou "hit" ou "stomp" (phase facultative de répétition) - nous appelons ces *types d'actions gestuelles variables*) et celles qui n'offrent pas la possibilité de maintenir ou de poursuivre l'action dans n'importe quel cas de production (par exemple, comme dans les actions gestuelles "beckon" ou "fling" qui passent toujours de la fin de l'action elle-même directement à la phase de récupération - nous appelons ces *types d'actions gestuelles variables*), comme dans les actions gestuelles "beckon" ou "fling" qui passent toujours de la fin de l'action elle-même directement à la phase de récupération - nous appelons ces *types d'actions gestuelles stables*). La figure 2 illustre la manière dont nous conceptualisons les différentes phases d'action des *types d'actions gestuelles A*) variables (illustrées par l'action gestuelle "atteindre") et B) stables (illustrées par l'action gestuelle "faire signe") dans notre schéma de codage. À l'instar de la préparation et de l'action de Kendon (2004), nous différencions une partie de l'action gestuelle qui devrait contenir l'information minimale nécessaire pour la distinguer de toute autre action gestuelle, que nous appelons l'*unité d'action minimale* (UAM). L'UAM de chaque type d'action gestuelle (codée une fois pour chaque instance de geste) commence au moment où le signaleur commence à faire un geste, c'est-à-dire qu'il déplace la partie du corps (*ici* le bras) hors de son état com- municativement "neutre" (ligne grise la plus claire dans la figure 2) et qu'il fait un geste qui n'a pas d'incidence sur le comportement du signaleur.

l'utilise pour effectuer le mouvement gestuel. Nous définissons un état com- municatif "neutre" lorsque la partie du corps employée dans le geste est au repos ou utilisée dans des actions non communicatives (par exemple, l'alimentation). L'UAM se termine lorsque l'action gestuelle est complètement en place (ligne grise la plus foncée sur la figure 2). Pour les types d'action gestuelle qui peuvent être maintenus en place (par exemple, "atteindre", figure 2A) ou répétés (rythmiquement) (par exemple, "frapper"), c'est-à-dire les *types d'action gestuelle variables*, l'action gestuelle elle-même ne se termine pas nécessairement avec la MAU mais peut se poursuivre jusqu'à ce que le signaleur décide d'arrêter d'afficher le geste et ramène la partie du corps concernée dans une position neutre (sur le plan de la communication) (par exemple, en rétractant son bras jusqu'à ce qu'il soit en contact avec le sol), en ramenant son bras dans une position de repos, ou en commençant à l'utiliser pour la locomotion, ou l'alimentation, etc.) ou, alternativement, l'utilise pour effectuer l'action gestuelle suivante. La phase facultative de maintien/répétition est indiquée par le remplissage gris de l'aérographe sur la figure 2A. Nous appelons *durée de l'action gestuelle* la durée qui s'écoule entre le début du premier mouvement à partir de la position neutre et la fin de l'action gestuelle. Nous faisons ici une distinction supplémentaire entre la durée de l'action gestuelle d'une *instance d'action gestuelle continue* (une instance de production de type d'action gestuelle variable où le signaleur a utilisé la phase facultative de maintien/répétition) et la durée de l'action gestuelle d'une *instance d'action gestuelle minimisée* (une instance de production de type d'action gestuelle variable où le signaleur a fait usage de la phase facultative de maintien/répétition).

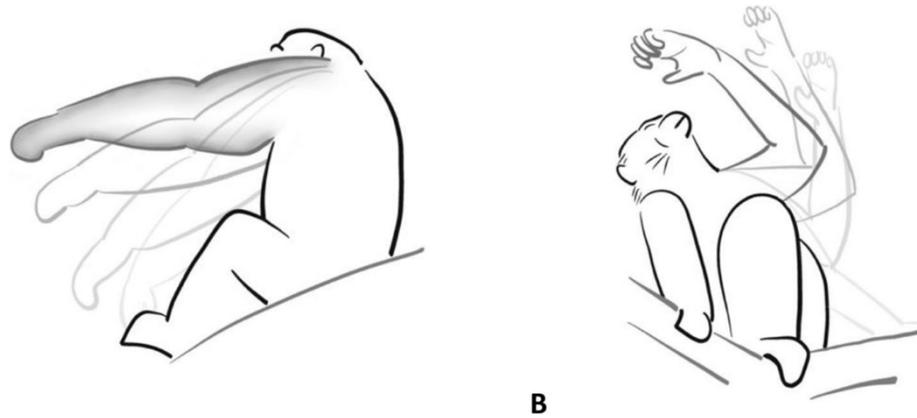


Fig. 2 Illustration des différentes phases de la course des types d'actions gestuelles variables et stables à l'aide des exemples "reach" et "beckon". *Remarque.* Nous partons du principe que le contenu informatif d'une action gestuelle peut ne pas être réparti uniformément sur toute la durée du geste, mais prendre plutôt la forme d'unités discrètes (par exemple, action et maintien), ce que nous illustrons à l'aide de deux exemples d'actions gestuelles. **Une "portée"** (un type d'action gestuelle variable) : l'unité d'action minimale (UAM) et l'action gestuelle (AG) commencent dès que le signaleur déplace son bras à partir de la position neutre (*ligne gris très clair*), c'est-à-dire qu'il commence à faire un geste. L'UAM se termine lorsque la phase de mouvement est achevée, c'est-à-dire lorsque la portée est dans son extension maximale vers le destinataire (*ligne gris foncé* ; durée de l'UAM = position neutre jusqu'à la fin de l'UAM). L'action gestuelle se poursuit jusqu'à ce que le signaleur commence à abaisser son bras, c'est-à-dire lorsque le geste n'est plus en place. Cette phase de maintien optionnelle entre la fin de l'action

MAU et la fin de l'AG est indiquée par le remplissage gris à l'aérographe (durée de l'action du geste = position neutre à la fin de l'AG, c'est-à-dire la fin de la phase de maintien/répétition). Nous annotons également le moment où le bras revient en position neutre afin de suivre le temps total investi dans la production du geste (durée du geste complet = position neutre à position neutre). **B "Beckon"** (un type d'action gestuelle stable) : la MAU et l'AG commencent dès que le signaleur bouge son bras de la position neutre (*ligne gris très clair*), c'est-à-dire qu'il commence à faire un geste. La MAU se termine lorsque la phase de mouvement est achevée (c'est-à-dire après l'action complète de "scooping beckon"). Comme il s'agit d'une action gestuelle stable, il n'y a pas de phase facultative de maintien/répétition, et l'action gestuelle se termine en même temps que la MAU (durée de la MAU/GA = position neutre à la fin de la MAU/GA). Comme pour la portée, nous annotons également le moment où le bras revient en position neutre (durée du geste = position neutre à position neutre).

production de type d'action gestuelle variable où le signaleur n'a pas utilisé la phase facultative de maintien/répétition ou l'instance appartient à un type d'action gestuelle stable). Par conséquent, dans les types d'action gestuelle stables, l'action gestuelle se termine typiquement en même temps que l'UAM (c'est-à-dire fin de l'action gestuelle = fin de l'UAM, car ils n'ont pas de phase de maintien/répétition : p. ex, Dans les types d'action gestuelle variables, l'action gestuelle peut ou non se terminer en même temps que l'UAM, en fonction de son utilisation spécifique dans la communication en cours (ESM : GOv1.0_Gesture_action_definitions.xlsx dans le matériel supplémentaire fournit des descriptions détaillées des actions gestuelles identifiées et de leurs phases d'action gestuelle respectives). Dans certains types d'actions gestuelles, la durée elle-même a été utilisée pour distinguer différentes catégories d'actions gestuelles (par exemple, toucher et toucher-long). Bien qu'il s'agisse de décisions *a priori* (basées sur des modèles d'utilisation établis dans des recherches antérieures), il s'agit également d'aspects qui peuvent être testés à l'avenir. Par exemple, en regroupant tous les cas d'actions tactiles et en traçant la durée des différents aspects des actions gestuelles afin d'explorer s'il existe des catégories apparentes en leur sein (par exemple, celles-ci pourraient être illustrées par une distribution bimodale ou multimodale des durées, par opposition à une distribution normale). Notre codage permet également de marquer les durées comme inconnues lorsqu'il n'est pas possible d'établir le début et/ou la fin d'une phase, ce qui permet d'exclure facilement ces cas des analyses.

Il est essentiel de savoir où et comment l'information est encodée dans les signaux gestuels pour étudier leur fonctionnement dans le système de communication. Dans notre codage, nous mesurons donc quatre points différents dans l'instance gestuelle : le début du geste (*Gesture start time*), la fin de l'unité d'action minimale (*MAU end time*), la fin de l'action gestuelle (*Gesture action end time*, dans le cas de types d'actions gestuelles stables, c'est toujours la même chose que la fin de la MAU ; dans le cas de types d'actions gestuelles variables, cela dépend de l'instance d'utilisation : soit à la fin de la MAU (lorsqu'aucune phase de maintien/répétition n'a été incluse), soit à la fin de la phase de maintien/répétition), et lorsque la partie du corps qui fait le geste est revenue à sa position neutre ou commence à produire une nouvelle action gestuelle (*heure de fin du geste*). À partir de ces données, nous pouvons calculer (après codage) la *durée totale du geste* (tous les mouvements impliqués dans la production de l'instance de geste, y compris sa récupération : de la position neutre à la position neutre), la *durée de l'action de geste* (le mouvement de l'action de geste et (éventuellement) son maintien en place : de la position neutre à l'heure de fin de l'action de geste) et la *durée de la MAU* (de la position neutre à l'heure de fin de la MAU), la partie de l'action de geste qui contient l'information permettant à l'instance de geste d'être "comprise" comme étant une action de geste ou un morphisme particulier.

Étant donné que l'UAM est la section du geste qui mène En plus d'être une unité reconnue par le destinataire, il est probablement soumis à des pressions physiologiques et sélectives pour une communication efficace, par exemple la compression (Heesen et al., 2019) ;

Safryghin et al., 2022), et nous nous attendons à ce qu'elle soit exprimée de manière assez cohérente d'une instance à l'autre. En revanche, si les grands singes intègrent l'utilisation d'une phase facultative de "maintien" dans leurs gestes, nous prévoyons une plus grande variation de la durée : la production d'un maintien est plus susceptible de dépendre de variables liées à l'instance spécifique de communication, telles que le temps de réponse du destinataire au geste ou la volonté du signaleur de s'investir dans un résultat particulier.

Exemple de travail 1 : Durée des gestes

Nous présentons ici un exemple concret où la durée des gestes est envisagée de différentes manières. Nous examinons les *durées MAU* de trois actions gestuelles courantes chez les gorilles de montagne ("beckon", "raise" et "reach" ; Grund et al., en préparation) et les comparons à leurs *actions gestuelles* respectives et à la *durée du geste complet* (voir Fig. 3). Nous nous attendions à ce que les durées des gestes soient plus variables pour les gestes "lever" et "atteindre" (car ils comportent une phase de maintien facultative) que pour le geste "faire signe", qui est plus stable. Pour évaluer la fiabilité des résultats, nous avons effectué des tests de corrélation intra-classe (ICC) à score unique (tests ICC2, modèle = "two-way", type = "agreement", *psych* package, v2.2.9, Revelle,

2022) sur les mesures "durée MAU" et "durée de l'action gestuelle" (en utilisant les temps G_start_T , MAU_end_T et GA_end_T , voir ESM_IRR_test.docx pour plus de détails) d'un petit sous-ensemble ($n = 40$ communications) de l'ensemble de données sur la latence ($n = 250$ communications, voir l'exemple de travail 2). La concordance globale entre les deux codeurs (CG, CH) était bonne pour la durée de l'UAM (ICC = 0,97, IC à 95 % de 0,95 à 0,99, $F(39, 39) = 73$, $p < 0,01$) et excellente pour la durée de l'action gestuelle (ICC = 0,99, IC à 95 % de 0,99 à 0,99, $F(39, 39) = 7796$, $p < 0,01$). Aucun biais de codage cohérent n'a été observé (voir ESM_IRR_test.docx dans ESM pour les graphiques montrant les différences de codage et plus de détails sur le test et l'ensemble de données utilisé).

Comme prévu, dans les actions gestuelles "lever" et "atteindre", il y a plus de variabilité dans la durée du geste entier (voir Fig. 3B : *durée de l'action gestuelle* et 3C : *durée du geste entier*) par rapport à leurs MAU respectives (voir Fig. 3A : *durée de l'unité d'action minimale*). Il est intéressant de noter que la durée de la phase de récupération (le temps écoulé entre la fin de l'action gestuelle et le retour au repos de la partie du corps) a beaucoup moins d'impact, comme le montre la différence plus faible entre la *durée de l'action gestuelle* et la *durée du geste complet*.

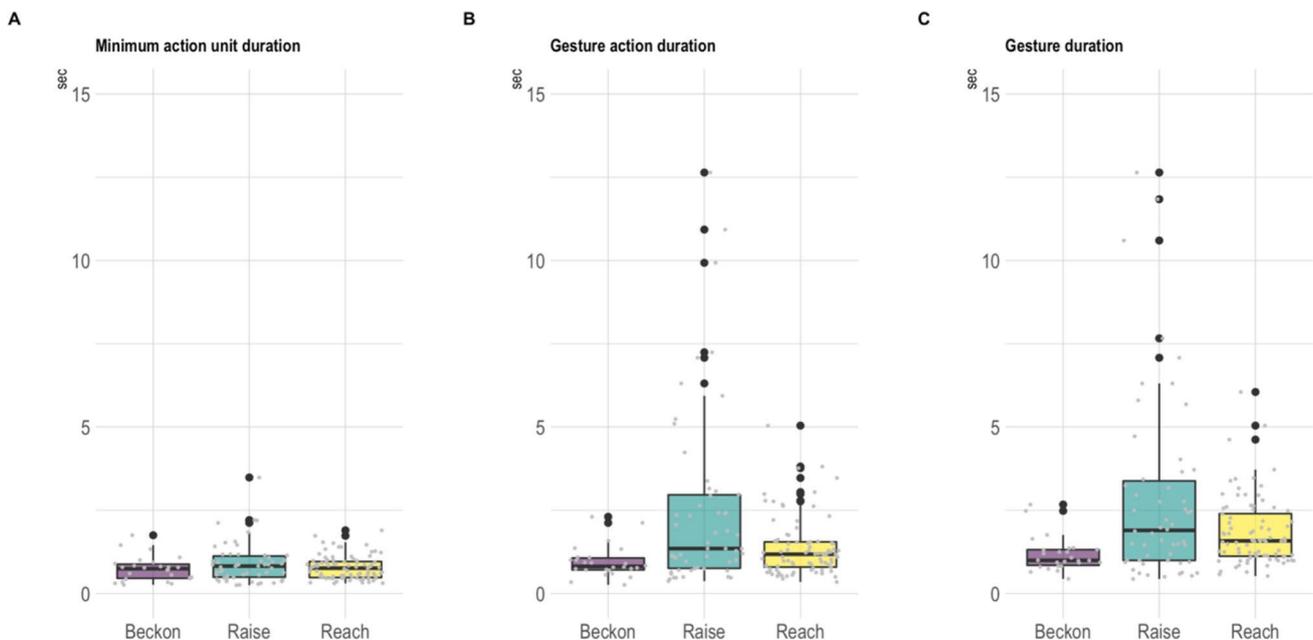


Fig. 3 Exemple de travail 1 : Différence de variabilité entre la durée de l'UAM, la durée de l'action du geste et la durée du geste complet. *Remarque.* Boxplots des actions gestuelles "beckon" ($n = 26$), "raise" ($n = 54$) et "reach" ($n = 85$) montrant la différence de variabilité de durée entre la phase MAU (*Minimum Action Unit*) d'une action gestuelle (*médiane de la durée MAU* pour "beckon" = 0,74 s (plage = 0,26-1,75 s) ; pour "raise" = 0,74 s (plage = 0,26-1,75 s) ; pour "reach" = 1,75 s (plage = 1,75 s)), 74 s (écart = 0,26-1,75 s) ; pour "raise" = 0,82 s (écart = 0,25-3,48 s) ; et pour "reach" = 0,76 s (écart = 0,30-1,90 s) ; *durée moyenne de l'UAM* pour "beckon" : 0,74 s (écart-type = 0,37) ; pour "lever" = 0,91 s (SD = 0,60) ; et pour "atteindre" = 0,79 s (SD = 0,35). **B** le temps nécessaire pour produire et maintenir l'action gestuelle (phase d'*action gestuelle* (AG) d'une action gestuelle : *Durée médiane de l'AG* pour "beckon" = 0,81 s (intervalle = 0,26-2,3 s) ; pour "raise" = 1,34 s (intervalle = 2,3 s) ; pour "raise" = 1,34 s (intervalle = 2,5 s) ; pour "reach" = 1,84 s (intervalle = 0,37-12,64 s) ; et pour "atteindre" = 1,18 s (intervalle = 0,35-5,04 s) ; *GA*

durée moyenne pour "beckon" : 0,95 s (écart-type = 0,48) ; pour "lever" = 2,55 s (écart-type = 2,75) ; et pour "atteindre" = 1,36 s (écart-type = 0,87)) et **C** incluant la phase de récupération du geste (*durée médiane du geste complet* pour "faire signe" = 1,00 s (écart-type = 0,44-2,67 s) ; pour "lever" = 1,90 s (écart-type = 0,44-2,67 s) ; pour "atteindre" = 1,90 s (écart-type = 2,75) ; et pour "atteindre" = 1,36 s (écart-type = 0,87)). = 0,44-12,64 s) ; et pour "atteindre" = 1,58 s (intervalle = 0,53-6,10 s) ; *durée moyenne du geste complet* pour "faire signe" : 1,17 s (écart-type = 0,51) ; pour "lever" = 1,58 s (écart = 0,53-6,10 s) ; la durée totale du geste est en moyenne de 2,87 s (SD = 2,82) ; et pour "reach" = 1,84 s (SD = 1,02)). Espèce = *Gorilla beringei beringei* ; $n = 165$ instances d'utilisation de gestes, $n = 26$ signaleurs, filtrés à partir de données de comportement gestuel de gorilles de montagne collectées sur quatre unités sociales (Mukiza, Oruzogo, Kyagurilo, Bitukura) dans le parc national de Bwindi Impénétrable, Ouganda, entre 2019-2022 - données dans ESM : Worked_example_1_data.csv).

semble se concentrer sur la phase de maintien. Comme prévu, ce schéma semble varier entre les types d'action gestuelle, avec une variation moindre de la durée entre l'action gestuelle et l'unité d'action miniaturée pour l'action gestuelle stable "faire signe".

Le fait que l'unité d'action minimale semble être constamment plus courte et moins variable en longueur que l'action gestuelle (avec des différences dans le degré de cette tendance entre les actions gestuelles) souligne l'importance de faire cette distinction en premier lieu. Il est important de noter qu'à la fin de l'unité d'action minimale, le signaleur a déjà communiqué au destinataire les informations essentielles concernant l'action gestuelle, ce qui a également des implications importantes sur la façon dont nous pourrions vouloir étudier les structures de séquence et calculer les latences de réponse (voir la section suivante).

Structure de la communication et latences : Interprétation des séquences de gestes et de l'attente de la réponse

De nombreuses communications peuvent être décrites comme des chaînes comportementales relativement simples - l'individu qui signale regarde le destinataire, fait un geste, marque une pause, le destinataire réagit par une réponse apparemment satisfaisante (ASO) et la communication s'achève. Cependant, certaines communications semblent structurellement plus complexes : par exemple, les signaleurs peuvent déployer plusieurs gestes (du même type ou de types différents) consécutivement (et/ou se chevauchant). Dans les communications comportant plusieurs gestes, un geste peut être directement suivi d'un autre avec peu ou pas de temps entre les deux ou, au contraire, le signaleur peut ne refaire un geste qu'après une pause plus longue, pendant laquelle il continue à surveiller le destinataire.

Une interprétation de ces structures est que dans le premier cas, les gestes sont produits comme des séquences rapides combinées indépendamment du comportement du destinataire, alors que dans le second, les gestes sont séparés par l'attente d'une réponse par le signaleur, les gestes suivants représentant la persistance après l'échec de la demande précédente (Liebal et al., 2004 ; Hobaïter & Byrne, 2011b ; McCarthy et al., 2013). Dans la plupart des approches, les chercheurs décident *a priori* de classer les gestes en séquences en fonction du temps qui les sépare. Dans certains cas, ces décisions dépendent d'une durée déterminée entre les gestes (Liebal et al., 2004 ; McCarthy et al., 2013), tandis que d'autres fondent les décisions de regroupement sur un cadre plus théorique, c'est-à-dire en combinant des durées basées sur la présence ou l'absence d'autres comportements dans l'intervalle (par exemple, Hobaïter & Byrne, 2011b). Par exemple, nous et d'autres avons utilisé un intervalle minimum (*a priori*) de 1 s entre le point final d'un geste et le début du suivant, pour indiquer que deux instances de gestes appartiennent à des séquences différentes, c'est-à-dire qu'il s'agit de deux instances de gestes avec une réponse en attente entre les deux (Hobaïter & Byrne, 2011b ; Luef & Liebal, 2012 ; Heesen et al., 2019).

Lorsque l'attente d'une réponse est considérée comme présente, les chercheurs interprètent le geste suivant comme une "persistance" (même type de geste) ou une "persistance avec élaboration" (nouveaux types de gestes) vers la réalisation du même objectif plutôt que comme la production d'une nouvelle communication (par exemple, Cartmill & Byrne, 2007 ; Leavens et al., 2005). La description de différents types de séquences gestuelles, par exemple l'ajout de gestes ultérieurs dans le cadre de la même communication, la persistance vers le même objectif ou une nouvelle communication, est utile pour distinguer les différents types de structures communicatives dans les gestes des grands singes. Cela signifie que l'application traditionnelle des règles est basée sur des intervalles de temps ou des indicateurs comportementaux qui ont été façonnés par l'étude d'un nombre relativement restreint de populations de chimpanzés et qui ne sont pas toujours appropriés pour étudier les gestes d'autres espèces de singes comme les gorilles de montagne ou les orangs-outans, sans parler des espèces non simiennes (Farrar et al., 2021). Même les bonobos, étroitement apparentés, semblent présenter des tendances différentes dans les temps d'attente des réponses (Fröhlich et al., 2016). Des limites rigides basées sur le temps ou le comportement pour décrire les structures gestuelles, telles que les séquences, peuvent donc être intrinsèquement problématiques lorsqu'on essaie d'étendre la recherche sur les gestes à d'autres espèces ou populations.

Il est également important de reconnaître que la durée de la l'intervalle entre deux gestes (et donc la possibilité d'atteindre une limite temporelle) dépend du moment de la phase d'action du geste que l'on considère comme un point de départ approprié pour le début de l'attente de la réponse. La figure 4 ci-dessous est une illustration schématique d'une communication (simple) dans laquelle un signaleur (A) fait deux gestes en direction du destinataire (B) et ce dernier répond par une OSA (c'est-à-dire un "but") après un certain laps de temps.

Si, par exemple, un geste d'extension est produit avec une extension du bras, puis que cette extension est maintenue en place pendant plusieurs secondes, on peut considérer que l'attente de la réponse commence a) lorsque le mouvement initial nécessaire pour produire l'action gestuelle est terminé (c'est-à-dire au point final de l'UAM, à la fin du mouvement d'extension - Fig. 4, couleur violette) ou b) lorsque l'action gestuelle est terminée et que le membre revient au repos ou commence à produire l'action gestuelle suivante (donc au point final de l'UAM), au point final de la MAU, à la fin du mouvement d'extension - Fig. 4, couleur violette) ou b) lorsque l'action gestuelle est terminée et que le membre est remis au repos ou commence à produire l'action gestuelle suivante (donc à la fin de la phase de maintien, indiquée à la Fig. 4 par la couleur grise, approche "traditionnelle"). Si nous exigeons que les signaleurs fassent preuve d'une pause d'au moins 1 s, ou d'un indicateur comportemental tel qu'une vérification visuelle pour marquer l'existence de l'attente de réponse, sa présence variera selon que nous mesurons à partir du point a) ou du point b). Les données illustrées dans l'exemple travaillé 1 confirment la suggestion selon laquelle le contenu informatif de certaines actions gestuelles peut ne pas être réparti uniformément sur toute la durée du geste et qu'il peut être approprié de considérer le point situé à la fin de la MAU (c'est-à-dire lorsque le signaleur a communiqué l'information principale dans la phase d'action) comme le début de l'attente d'une réponse.

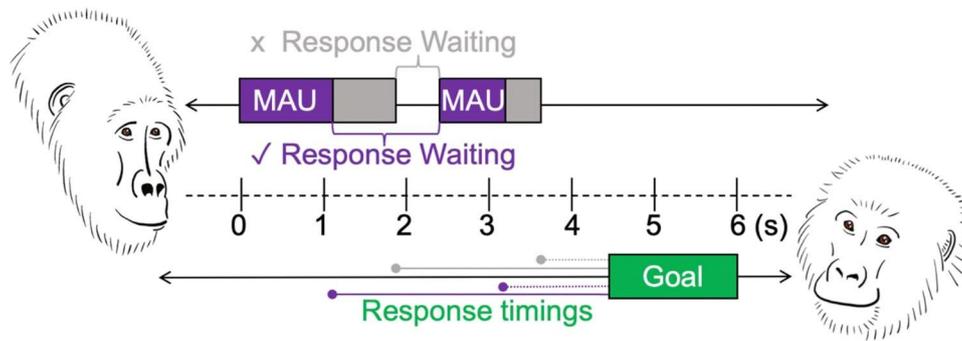


Fig. 4 Visualisation des différentes façons de mesurer les latences de réponse et les intervalles entre les gestes à l'aide du système de codage GesturalOrigins. *Remarque.* Chaque communication peut être considérée comme une chaîne de comportements sur une ligne temporelle qui commence par le geste d'un individu vers un autre et se termine par un résultat comportemental particulier (typiquement, soit l'objectif de l'individu signalant (ASO), soit l'échec de la communication). Dans cet exemple de communication gestuelle simple, l'individu A (le signaleur) à gauche fait deux gestes et l'individu B (le destinataire) répond par un résultat apparemment satisfaisant (le "but" de la communication) après un certain temps. Si nous considérons que l'intervalle entre les gestes commence à la fin de la durée de l'action du geste et se termine au début du geste suivant, il est inférieur à 1 s et ne correspond pas à la limite temporelle pour l'attente d'une réponse, et les deux gestes seraient considérés comme faisant partie de la même séquence. Cependant, lors de l'étude des intervalles entre les gestes et des temps d'attente de la réponse, les

La fin de la MAU peut être particulièrement intéressante, car elle correspond à peu près au moment où toutes les informations nécessaires pour que l'action gestuelle soit comprise comme un cas de cette action gestuelle particulière devraient être en place. Si nous considérons que l'intervalle inter-gestes commence à la fin de la MAU et se termine au début du geste suivant, l'intervalle de temps dans cet exemple est alors > 1 s et correspondrait à la limite temporelle pour l'attente d'une réponse, et les deux instances de gestes ne feraient pas partie de la même séquence. Outre les intervalles entre les gestes, le système de codage offre une grande flexibilité dans le calcul des latences de réponse du destinataire (voir l'exemple de travail 2). Certains d'entre eux sont indiqués par des lignes violettes et grises partant de différents points de fin de geste jusqu'au *résultat* (boîte de but verte). Notez que par souci de clarté et compte tenu du fait que la différence entre la durée de l'action du geste et la durée totale du geste semble cohérente et faible (voir l'exemple de travail 1), aucune distinction n'est faite entre ces deux durées dans le graphique.

Qu'en est-il des points finaux de l'attente d'une réponse ? Du point de vue de la communication dans son ensemble (et du fait que nous supposons un objectif global), l'attente d'une réponse ne peut être considérée comme terminée que lorsque le destinataire modifie son comportement de la manière souhaitée et que l'objectif visé par le signaleur (c'est-à-dire la raison pour laquelle il a fait un geste) a été atteint. Avec la variable *temporelle Outcome*, nous marquons le moment où l'objectif est atteint et où la communication est terminée. Lorsque l'*issue* ne peut être déterminée (par exemple, en cas d'échec de la communication), elle est marquée comme "inconnue" (mais voir ESM : GOv1.0_Elan_controlled_vocabulary.xlsx - Sheet 'Goal' pour une liste des objectifs identifiés du signaleur). La variante du *temps de sortie*, associée à l'*UAM* et aux temps d'*action* des instances de geste, nous permet de calculer différents temps d'attente de réponse, par exemple, des temps d'attente du point de vue de l'ensemble de la communication qui peuvent ne pas s'arrêter au début du geste suivant mais qui s'étendent de l'instance de geste actuelle à la fin réelle de la communication, ou des temps d'attente qui, par exemple, ne considèrent que le dernier geste de la communication comme pertinent lors du calcul des latences de réponse du destinataire (comme dans l'exemple travaillé n° 2). Comme le signaleur tente de modifier le comportement du destinataire tout au long de la communication et qu'il est susceptible d'adapter ses gestes au comportement du destinataire (ou de négocier, par des échanges gestuels avec le destinataire), il peut également être intéressant d'explorer les réponses du destinataire (latences) avant tout résultat communicatif (satisfaisant). C'est pourquoi nous avons également codé les temps suivants des destinataires sur la page d'accueil.

niveau de la communication : le destinataire a-t-il répondu par un geste (*Gesture recipient*, ESM : GOv1.0_Protocol - section 2.13) ou une vocalisation (*Vocalisation recipient*, ESM : GOv1.0_Protocol - section 2.14) à un moment ou à un autre de la communication ? Le destinataire a-t-il réagi à la gestuelle d'une manière ou d'une autre (*changement de comportement 1 et 2*, ESM : GOv1.0_Protocol - sections 2.17 et 2.18) avant la fin de la communication ?

Avec le schéma GesturalOrigins, les chercheurs peuvent calculer les intervalles entre les gestes et les latences de réponse du destinataire à partir de différents points de fin de geste (par exemple, la *fin de l'UAM* ou la *fin de l'action du geste*) de manière flexible et étudier leurs effets sur la dynamique d'interaction dans la communication. Notre prochain exemple pratique montre pourquoi le codage à haute résolution de données sensibles au temps et une approche axée sur les données pour l'étude de la structure des séquences et des latences peuvent être cruciaux pour étendre l'étude des gestes à de nouvelles espèces.

Exemple pratique 2 : Quelle est la durée d'attente d'une réponse ?

En utilisant les données gestuelles de chimpanzés d'Afrique de l'Est (deux communautés : Sonso et Waibira, forêt de Budongo, Ouganda) et de gorilles de montagne (quatre unités sociales : Mukiza, Oruzogo, Kyagurilo et Bitukura, parc national de Bwindi Impénétrable, Ouganda), nous avons étudié deux mesures différentes du temps dans 957 communications réussies (gorilles de montagne :

$n = 250$ communications, $n = 37$ signaux ; chimpanzés : $n = 707$ communications, $n = 115$ signaleurs). Nous avons calculé la latence à la fois a) de la fin de la MAU au résultat (*latence MAU.Outcome*, c'est-à-dire du moment où la phase d'action du geste a été achevée jusqu'à ce que le but soit atteint, voir ligne pointillée violette dans la Fig. 4) et ensuite b) de la fin de toute l'action du geste au résultat (*latence GA.Outcome*, c'est-à-dire du moment où le signaleur a cessé de montrer le geste jusqu'à ce que le but soit atteint, voir ligne pointillée grise dans la Fig. 4). Lorsque la communication comportait plusieurs gestes, nous n'avons pris en compte que l'UAM ou l'AG du dernier geste (celui qui est le plus proche du résultat). La figure 5, tracés A et B, montre les données pour un seul geste ("présent") couramment utilisé chez les deux espèces pour initier le "toiletage" ($n = 281$ communications réussies, les données (ESM : Worked_example_2.1_data.csv) n'incluent que les pré-sentations pour le toiletage afin de contrôler les variations possibles dues à l'objectif de l'appelant).

Les données présentées dans les graphiques indiquent une tendance générale des gorilles de montagne à prendre plus de temps que les chimpanzés pour répondre aux cadeaux de toiletage (gorille de montagne *MAU.Résultat*

latence : plage = - 0,72-78,1 s, médiane = 2,45 s, moyenne = 4,9 s (écart-type = 9,48), $n = 104$ communications ; chimpanzé d'Afrique de l'Est *MAU.Résultat latence* : plage = - 1,5-23,6 s, médiane = - 2,45 s, moyenne = 4,9 s (écart-type = 9,48), $n = 104$ communications ; chimpanzé d'Afrique de l'Est *MAU*.

0,09 s, moyenne = 1,52 s (écart-type = 2,56), $n = 177$ communications). Pour voir si cette différence apparente entre les espèces dans la gestuelle est observée plus globalement, nous avons reporté les données pour toutes les actions gestuelles (et tous les buts à l'exception de "jouer" ; Fig. 6A-D ; données en ESM : Worked_example_2.2_data.csv).

Bien que préliminaires, les figures 5 et 6 suggèrent ensemble des différences potentiellement notables entre les espèces dans la dynamique interactionnelle entre les chimpanzés et les gorilles de montagne. Le plus évident est une réactivité plus lente du destinataire aux gestes chez les gorilles de montagne (par exemple, les valeurs moyennes plus élevées de la latence *MAU.Outcome* chez MG (moyenne = 4,73 s (SD = 8,55), $n = 250$) par rapport à EAC (moyenne = 4,73 s (SD = 8,55), $n = 250$) par rapport à EAC (moyenne = 4,73 s (SD = 8,55), $n = 250$).

= 2,08 s (SD = 3,46), $n = 707$, voir Fig. 6). Deuxièmement, alors que les deux espèces montrent une différence relative claire (et assez similaire) entre la *durée de l'action gestuelle* et la *durée de la MAU*, les gorilles de montagne montrent une variabilité potentiellement plus grande de la latence lorsqu'elle est mesurée à partir de la fin de la MAU ou de la fin de l'action gestuelle (divergence dans les valeurs entre *MAU.Outcome*

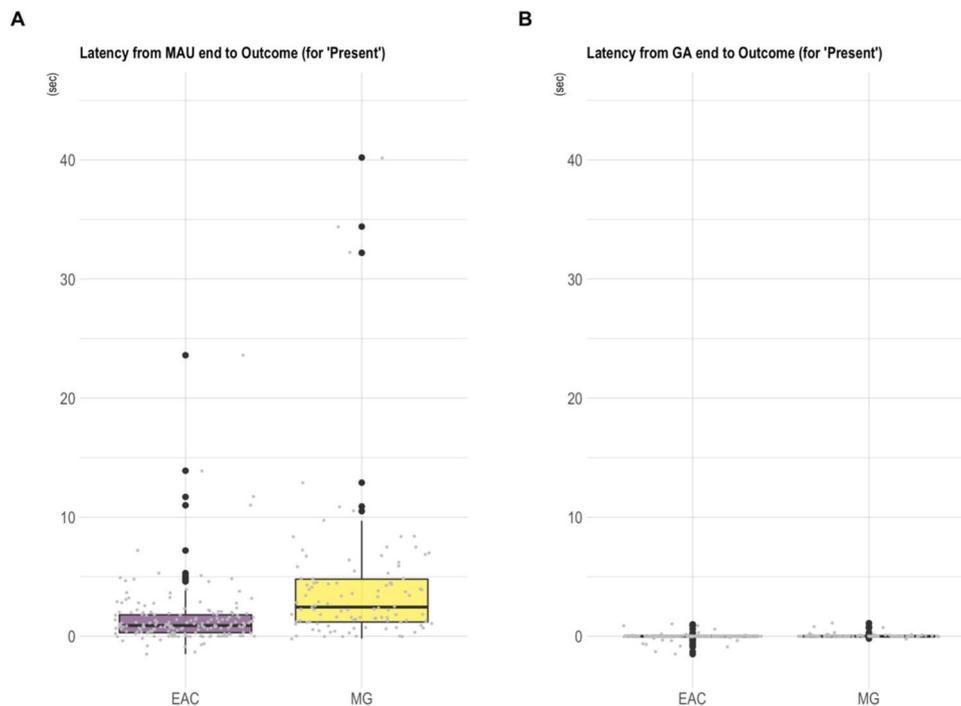


Fig. 5 Exemple de travail 2.1 : Temps de latence des gorilles de montagne et des chimpanzés pour répondre (comportementalement) à la demande de toiletage d'un signaleur en utilisant l'action gestuelle "présent" ($n = 281$ communications réussies). *Remarque*. Diagrammes en boîte montrant les temps d'attente de réponse (latences pour répondre) chez les chimpanzés d'Afrique de l'Est (EAC ; $n = 177$ communications, $n = 70$ signaleurs) et les gorilles de montagne (MG ; $n = 104$ communications, $n = 27$ signaleurs) pour l'action gestuelle "présent" et le résultat "toiletage" (total : $n = 281$ communications) en considérant soit **A** le point final de l'unité d'action minimale (MAU), soit **B** le point final de l'action gestuelle (GA) comme le début de l'attente de la réponse (voir Fig. 4 pour une visualisation conceptuelle des différentes mesures de latence de réponse). Les gorilles de montagne semblent être plus lents à répondre aux

Les chimpanzés de l'Afrique de l'Est ont plus de demandes de toiletage que les chimpanzés de l'Afrique de l'Est. Notez qu'une valeur unique (la valeur maximale de 78,1 s pour les gorilles de montagne) n'est pas représentée visuellement sur le graphique (bien qu'elle ne soit pas omise dans l'échelle) pour des raisons de meilleure résolution. Comme l'action gestuelle "présente" a la caractéristique d'être maintenue en place jusqu'à ce que l'objectif soit atteint (dans les communications réussies), il n'est pas surprenant qu'il y ait peu de variation dans la latence entre l'action gestuelle et le résultat (graphique B), qui est proche de 0 chez les deux espèces (MG *GA.Outcome* latency range = - 0.2- 1.1 ; EAC *GA.Outcome* latency range = - 1.5-1.0 s). Les latences négatives résultent des cas où le destinataire a déjà répondu au geste avant que la phase d'action respective (MAU et/ou GA) ne soit terminée

et la latence de *GA.Outcome* dans EAC (latence de MAU Outcome moyenne = 2,08 s (SD = 3,46) vs. *GA.Outcome* latency moyenne = 2,08 s (SD = 3,46) vs. *GA.Outcome* latency moyenne = 2,08 s (SD = 3,46)) dans EAC. = 0,71 s (SD = 2,25) ; $n = 707$; et en MG (MAU Outcome latency mean = 4,73 s (SD = 8,55) vs. *GA.Outcome* latency mean = 1,58 s (SD = 3,56) ; $n = 250$).

Les données présentées ici (gorilles de montagne $n = 250$ communications ; chimpanzés $n = 707$ communications) ont été codées par trois chercheurs différents (AS, CG, GB) et les différences entre espèces pourraient être le résultat de biais de codage (entre CG et AS/GB). Pour mesurer la fiabilité des résultats, CG (codage du gorille de montagne) a codé des sous-ensembles de l'ensemble de données sur les chimpanzés (*Sonso* : 25/555 communications codées par AS ; *Waibira* : 20/152 communications codées par GB) et a effectué des tests de corrélation intra-classe (ICC) à score unique (tests ICC2, modèle = "two-way", type = "agreement", *psych* package, v2.2.9, Revelle, 2022) sur les durées *MAU.Outcome* latency et *GA.Outcome* latency ainsi que sur les mesures de temps utilisées pour les calculer (voir ESM : IRR_test pour plus de détails sur les tests ICC2 portant sur MAU_end_T, GA_end_T et Outcome_T séparément). Dans l'ensemble, la concordance est bonne

entre AS et CG (MAU.Outcome latency ICC = .92, 95% CI de .79 à .97, $F(24, 24) = 29$, $p < 0.01$) ; *GA.Outcome* latency ICC = 0,95, IC 95 % de 0,88 à 0,98, $F(24, 24) = 50$, $p < 0,01$) et entre GB et CG (MAU.Outcome latency ICC = .95, 95% CI de .85 à .98, $F(19, 19) = 45$, $p < 0,01$) ; Les résultats de l'étude ont été comparés à ceux de l'étude sur la latence de l'AG ICC = 0,98, IC à 95 % de 0,95 à 0,99, $F(19, 19) = 88$, $p < 0,01$) dans le codage des variables de latence. Dans les graphiques des écarts par rapport à la moyenne pour les variables de latence MAU.Outcome et *GA.Outcome*, l'inspection visuelle a suggéré un biais possible en faveur de latences plus longues dans CG par rapport à AS (différence de codage moyenne MAU.Outcome latency = 0,48 ; différence moyenne de codage *GA.Outcome* latency = 0,38, très probablement due à des différences cohérentes dans le codage de la variable Outcome time, voir ESM : IRR_test.docx, Figure ESM.2 pour plus de détails). Pour déterminer si les différences entre les espèces pouvaient s'expliquer par un biais du codeur, nous avons ajouté la différence moyenne de codage entre les deux codeurs (CG et AS) à toutes les valeurs de latence de l'AS. Nous avons continué à trouver une différence cohérente dans les latences entre les espèces (voir ESM : IRR_test.docx, Figure ESM.4).

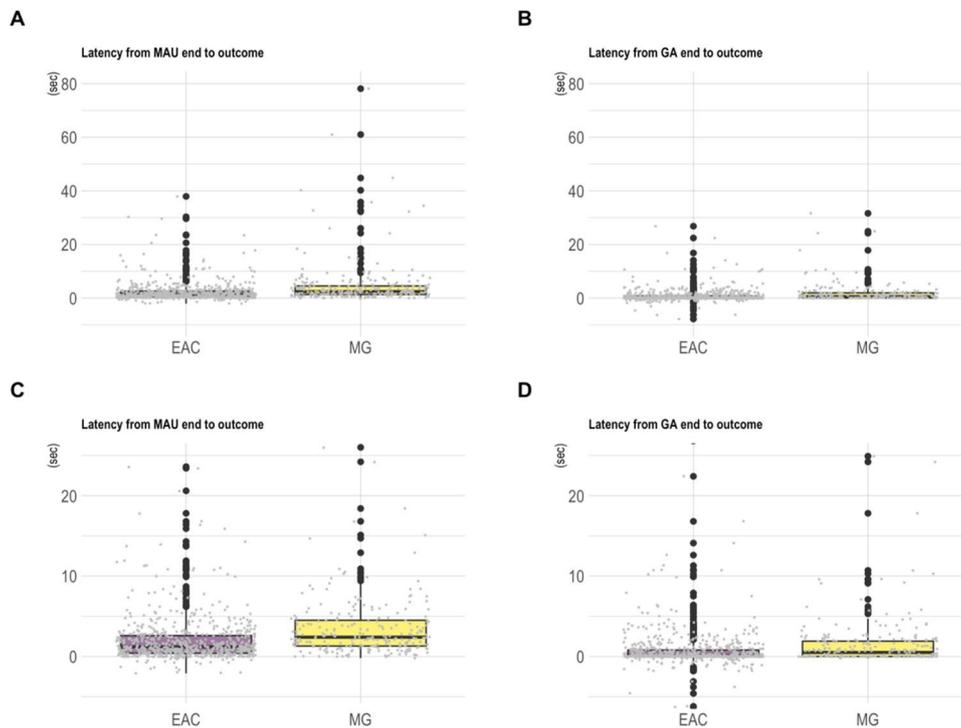


Fig. 6 Exemple de travail 2.2 : Temps de réponse (comportementale) du gorille de montagne et du chimpanzé aux gestes d'un signaleur (communications réussies, tous les objectifs sauf "jouer"). *Note.* Diagrammes en boîte montrant les temps d'attente de réponse (latences pour répondre) chez les chimpanzés d'Afrique de l'Est (EAC) et les gorilles de montagne (MG) dans 958 communications réussies (MG : $n = 250$ communications, $n = 37$ signaleurs ; EAC : $n = 707$ communications, $n = 115$ signaleurs - y compris toutes les actions gestuelles, à l'exception de l'objectif "jeu" - données dans ESM : Worked_example_2.2_data.csv) en considérant soit le point final de l'UAM (A et C), soit le point final de l'action gestuelle (B et D) comme début de l'attente de la réponse. Le graphique A comprend la gamme complète des valeurs de latence observées pour la fin de l'UAM jusqu'au résultat (gamme :

- 2,1-78,1) et le graphique B la gamme complète des valeurs de latence pour l'action gestuelle de la fin à l'issue (gamme : - 7,8-31,6), tandis que les graphiques C et D ne montrent que les latences dont les valeurs sont comprises entre - 5 et 25 s (pour une meilleure résolution de l'endroit où se trouve la majorité des données, sans pour autant omettre les valeurs les plus extrêmes de la mise à l'échelle). Les données suggèrent que les gorilles de montagne mettent plus de temps à répondre aux demandes gestuelles que les chimpanzés si l'on considère le point final de l'UAM comme le début de l'attente de la réponse et si l'on considère le point final de l'AG comme le début de l'attente de la réponse. Comme dans l'exemple 2.1, les latences négatives résultent de cas où le destinataire a déjà répondu au geste avant que la phase d'action respective (MAU et/ou GA) ne soit achevée

Les exemples travaillés que nous présentons ici servent à illustrer les types d'analyses potentielles que notre système de codage offre. Une analyse appropriée des différences interspécifiques, par exemple en ce qui concerne les temps de réponse, devrait prendre soigneusement en compte d'autres sources de variation, par exemple les différences au niveau de l'individu et/ou du groupe. Un point de notre approche que nous souhaitons souligner est que - comme il s'agit d'un nouveau concept - nous avons défini les UAM pour chaque action gestuelle *a priori* en utilisant les informations minimales dont nous avons besoin en tant qu'observateurs humains pour pouvoir distinguer les actions gestuelles (voir les définitions dans la feuille GOv1.0_Gesture_action_definitions.xlsx). Les parties du mouvement gestuel que nous considérons comme cruciales ne reflètent pas toujours l'information minimale nécessaire aux singes pour distinguer une action gestuelle d'une autre. Ces définitions gagneront à être testées à l'avenir avec des données codées détaillées. Un test potentiel de leur validité pour une action gestuelle particulière pourrait consister à utiliser les latences de réponse. Dans ce cas, lorsque les UAM d'une action gestuelle semblent être comprises "trop tôt" dans toutes les utilisations (c'est-à-dire que le temps de réponse est plus court que l'UAM - représenté par des latences négatives de l'UAM par rapport au résultat), indépendamment de l'objectif ou des individus en interaction, il peut être approprié de reconsidérer la définition de cette UAM. D'un autre côté, si nous ne trouvons pas de latences MAU négatives cohérentes par rapport au résultat, nous pouvons être plus sûrs que nous n'avons pas surestimé la partie du mouvement gestuel qui encode des informations importantes.

Combinaisons multisignaux : intégration des expressions faciales et des vocalisations

La recherche sur les gestes a été critiquée pour avoir négligé d'autres éléments communicatifs tels que les vocalisations ou les expressions faciales, généralement en les excluant de la collecte de données. Étudier les gestes de manière isolée, plutôt que de considérer la communication dans son ensemble, limite notre interprétation des capacités communicatives d'autres espèces (par exemple, Slocombe et al., 2011 ; Genty, 2019 ; Liebal et al., 2022). En même temps, le codage est une entreprise qui prend énormément de temps et il y a toujours un compromis entre l'effort de codage, la fiabilité du codage et le rendement de l'information. Les approches fiables du codage holistique peuvent bénéficier de l'utilisation de plusieurs caméras synchronisées pour permettre de capturer le signaleur (et idéalement aussi le destinataire) sous plusieurs angles, garantissant la détection de toute la gamme des différents types de signaux. Mais cela augmente considérablement l'investissement dans la collecte de données vidéo et dans le codage ultérieur des données. Dans la pratique, bien que le codage holistique soit idéal dans la mesure du possible, des sacrifices sont nécessaires. Par exemple : dans nos travaux actuels, nous n'incorporons pas de codage continu des signaux vocaux et faciaux, et nous les marquons uniquement lorsqu'ils sont associés à des gestes. De cette manière, chaque fois qu'il y a une instance de geste, les signaux non gestuels qui se chevauchent dans le temps avec les gestes sont capturés et leur fonction communicative potentielle peut être étudiée. Cependant, la fonction GesturalOrigins

est adaptable aux besoins de projets spécifiques et comprend des niveaux qui permettent aux utilisateurs de coder les signaux vocaux et faciaux en continu, tout au long de la communication. Il est également facile d'ajouter de nouveaux niveaux qui précisent le codage détaillé de signaux supplémentaires ou qui intègrent d'autres logiciels pour des questions de recherche spécifiques. Par exemple : Elan dispose d'une option intégrée pour utiliser "Praat" (Boersma & Weenink, 2022), un logiciel libre pour l'analyse acoustique.

Influence potentielle des facteurs socio-écologiques sur la gestuelle

On comprend de mieux en mieux l'importance de la pragmatique dans la communication des primates (Seyfarth & Cheney, 2018 ; Arnold & Bar-On, 2020), par exemple, des travaux récents sur la gestuelle des singes ont montré que l'ambiguïté apparente dans la signification de certains gestes des bonobos peut être résolue en incluant des informations sur le contexte comportemental dans lequel ils sont produits (Graham et al., 2020). Pour chaque communication, nous codons le contexte comportemental de base des interactants avant le début des gestes et après la fin de la communication (par exemple, "repos", "toiletage", "sexe" ; voir ESM : GOv1.0_Controlled_vocabulary.xlsx - Sheet 'Context' pour une liste de tous les contextes comportementaux). De même, l'utilisation des gestes par les individus peut varier en fonction de leur socialité et de leurs relations sociales (par exemple, Bard et al., 2014 ; Pika & Fröhlich, 2018 ; Fröhlich et al., 2022 ; Hobaiter & Byrne, 2011b). La possibilité d'inclure des informations sociales et environnementales plus riches sur la production de gestes semble donc importante pour affiner notre étude de son utilisation. La variation de l'histoire de vie et du contexte socio-écologique de la production peut façonner l'utilisation des gestes différemment entre les individus, les communautés et les espèces (Graham et al., 2022). De nombreux aspects, tels que la taille de l'individu, de l'espèce, de la communauté, la cohésion, l'âge, le sexe, le rang social ou la connectivité sont indépendants du codage immédiat d'une instance de geste et n'ont donc pas besoin d'être codés directement avec chaque instance. Cependant, certains aspects de la production de gestes sont spécifiques à une communication particulière et dans GesturalOrigins, nous incluons la possibilité de coder des variables telles que l'emplacement du signaleur et du destinataire (par exemple, arbre, sol ; voir ESM : GOv1.0_Protocol - sections 2.47, 2.51) pendant le geste, ou la visibilité (voir ESM : GOv1.0_Protocol - section 2.53) et la distance spatiale (voir ESM : GOv1.0_Protocol - section 2.40) qui les séparent.

Discussion et perspectives

L'étude des gestes, comme de nombreux aspects du comportement, bénéficie considérablement de l'utilisation de méthodes reproductibles et transparentes. L'un des moyens d'y parvenir est d'utiliser des cadres de codage communs qui permettent une comparaison similaire des recherches au sein des groupes de recherche et entre eux. Toutefois, dans certains cas, leur adoption peut s'avérer difficile lorsque

les différents groupes de recherche ont des approches bien établies mais variables pour coder et décrire leurs données. Cette variation n'est pas problématique en soi, mais elle pose un problème de comparabilité et de reproductibilité. Avec le cadre GesturalOrigins, nous visons à fournir une approche ascendante du codage des données qui permette aux chercheurs d'extraire les caractéristiques de l'inter- est de manières très diverses, en construisant des unités gestuelles qui varient en longueur ou en construction et qui sont déployées avec des critères différents. De nombreux critères établis utilisés dans la recherche gestuelle existante, par exemple l'utilisation d'un intervalle de temps particulier pour discriminer les séquences, peuvent encore être extraits. Mais surtout, il est possible de les faire varier et, par conséquent, de les comparer, de les tester et de les valider plus facilement. Pour ce faire, il faut que les méthodes, y compris les définitions des variables (par exemple, les objectifs, les actions gestuelles, etc.), soient décrites de manière suffisamment détaillée au moment de la publication (Rodrigues et al., 2021).

L'avantage de ces nouveaux outils méthodologiques est qu'ils permettent non seulement d'améliorer la reproductibilité, mais aussi d'explorer de nouvelles questions. En utilisant un logiciel d'annotation vidéo, nous pouvons examiner les unités gestuelles avec des détails de plus en plus fins - par exemple, en distinguant les parties d'un geste qui sont nécessaires à la production dans tous les cas (comme la préparation et le trait d'action) de celles qui sont facultatives et peuvent varier avec chaque instance de production (comme le maintien, la répétition ou la récupération). Dans nos exemples pratiques, nous montrons que certains aspects de la durée des gestes varient de manière non triviale. Non seulement les gestes du même type sont plus longs lorsque nous considérons des phases de production supplémentaires, mais ils présentent une variation substantielle de la durée au sein d'une phase, ce qui suggère que ces phases représentent des ajouts facultatifs. Une meilleure compréhension de la manière dont l'information est encodée et transmise dans la signalisation gestuelle peut s'avérer cruciale pour une variété de questions de recherche. Par exemple, des études récentes sur la loi de compression de Zipf - qui postule que les signaux les plus fréquemment utilisés sont plus courts (et qui est répandue dans les systèmes de communication humains et non humains ; Zipf, 1949 ; Ferrer-i-Cancho et al., 2022 ; Favaro et al., 2020) - ne montre qu'une évidence limitée dans les gestes des chimpanzés (Heesen et al., 2019). Une explication possible est que, jusqu'à présent, la mesure des durées complètes des gestes peut masquer son effet (Safryghin et al., 2022). La compression d'un signal dans le répertoire d'un individu ou d'une espèce devrait agir sur les parties du signal nécessaires à chaque instance de sa production (dans notre cadre, l'UAM), mais ce schéma peut être masqué si nous considérons des éléments facultatifs de leur production qui varient dans chaque instance de communication - comme une prise prolongée d'une portée lors de la demande d'un article alimentaire particulièrement précieux, ou la répétition prolongée d'un coup d'action du geste de *secouer l'objet* lors d'une sollicitation sexuelle importante. Ainsi, en utilisant des approches ascendantes telles que GesturalOrigins, nous pouvons extraire de manière flexible les minutages de différents points de l'action gestuelle (par exemple, à la fin de la MAU ainsi qu'à la fin de la "prise") et tester comment ils peuvent avoir un impact sur la structure de la séquence et façonner la communication gestuelle.

L'exploration d'hypothèses évolutives sur la trajectoire de la gestuelle dans la lignée des hominidés nécessite des données combinées sur diverses espèces et populations de grands singes. Pour ce faire, il faut non seulement des ensembles de données vraiment vastes, mais aussi que notre approche du codage évite un biais centré sur l'espèce qui pourrait ne pas décrire avec précision la gestuelle d'autres espèces de singes. À l'heure actuelle, environ 75 % des études sur les gestes des grands singes ont été menées sur des chimpanzés (Rodrigues et al., 2021). Les premières comparaisons suggèrent déjà des différences entre les espèces dans l'expression des formes gestuelles (par exemple, l'utilisation des membres entre les chimpanzés et les orangs-outans ; Knox et al., 2019) et dans la réactivité (chimpanzés et bonobos ; Fröhlich et al., 2016). La durée des intervalles de temps qui reflètent les connexions ou les distinctions dans les chaînes de comportement, par exemple l'attente de la réponse ou la latence pour répondre, peut également être façonnée par la socio-écologie de l'espèce. Dans notre deuxième exemple pratique, nous montrons que les gorilles de montagne présentent généralement des temps de latence plus longs pour répondre aux gestes que les chimpanzés. Ainsi, les comparaisons entre espèces utilisant un intervalle fixe peuvent ne pas fournir une comparaison similaire des comportements, et il faut être prudent dans l'application des lignes de base établies chez les chimpanzés, afin d'éviter de mal caractériser la communication gestuelle chez d'autres grands singes. Bien entendu, notre système de codage présente également des limites. Par exemple, les catégories comportementales doivent être établies et définies avant le codage, et des biais peuvent être introduits à ce niveau. Les options disponibles pour une variable particulière (par exemple, la liste des contextes, des objectifs ou des types d'action gestuelle) peuvent être plus appropriées pour une espèce que pour une autre (par exemple, parce que cette dernière n'a pas encore été étudiée de manière approfondie) ou plus détaillées pour les comportements fréquents que pour les comportements plus rares. Toutefois, si le recodage a un coût, il peut être ajusté au fur et à mesure que l'on acquiert de l'expérience sur une espèce ou dans un contexte donné. Les exemples vidéo en libre accès (comme le dictionnaire des grands singes) peuvent contribuer à améliorer la reproductibilité des études, car les chercheurs peuvent avoir accès à des exemples vidéo d'actions gestuelles qui peuvent être utilisés pour déterminer si une action particulière est présente dans un ensemble de données ou une espèce spécifique.

D'autres outils méthodologiques nouveaux et passionnants sont en cours d'élaboration.

horizon, ce qui peut augmenter l'efficacité des outils actuels tels que notre schéma de codage. Par exemple, nous commençons à déléguer certains des aspects du codage vidéo qui prennent le plus de temps, comme la détection des espèces (Beery et al., 2019), des individus (Schofield et al., 2019) et des comportements (Bain et al., 2021) à des modèles d'apprentissage automatique. Bientôt, il sera peut-être possible de déléguer la détection des interactions sociales (qui peuvent contenir des gestes) dans de grands corpus vidéo tels que le Great Ape Video Ark (Hobaiter et al., 2021 ; Wiltshire et al., en cours de révision) à des approches automatisées similaires, et même le comportement gestuel lui-même et les instances gestuelles codées soigneusement par la main et l'œil pourront éventuellement être détectés avec précision à l'aide d'approches d'apprentissage automatique. Même si tous les aspects du codage ne seront probablement pas facilement automatisés, pour les variables où cela est possible, de grands ensembles d'entraînement pré-annotés avec des données horodatées codées avec précision seront essentiels pour l'entraînement des modèles.

Ainsi, le codage vidéo basé sur des modèles peut également faciliter des étapes méthodologiques importantes dans le domaine de la recherche sur les gestes.

Remerciements Cette recherche a été financée par le 8e programme-cadre de l'Union européenne, Horizon 2020, dans le cadre de la convention de subvention n° 802719. Nous remercions Martha M. Robbins pour son soutien dans la collecte de données sur les gorilles de montagne. Nous remercions le personnel de terrain de Waibira, Sonso et Bwindi, ainsi que tout le personnel de la Budongo Conservation Field Station et de l'Institute for Tropical Forest Conservation pour leur soutien. Nous remercions le Conseil National Ougandais pour la Science et la Technologie, le Bureau du Président, l'Autorité de la Faune Ougandaise, et l'Autorité Forestière Nationale pour la permission de mener cette recherche et la Société Royale Zoologique d'Ecosse pour son soutien à la Station de Terrain de Conservation de Budongo.

Disponibilité du code Tout le code utilisé pour ce manuscrit est disponible sur GitHub (https://github.com/CharlotteGrund/Gestural_Origins_Coding-methods_paper).

Financement Cette recherche a été financée par le 8e programme-cadre de l'Union européenne, Horizon 2020, dans le cadre de la convention de subvention n° 802719.

Disponibilité des données Toutes les données et tous les documents mentionnés dans cet article sont disponibles sur GitHub (https://github.com/CharlotteGrund/Gestural_Origins_Coding-methods_paper) et/ou inclus dans cet article publié (et ses fichiers d'informations supplémentaires).

Déclarations

Approbation éthique L'approbation éthique pour la collecte des données vidéo originales utilisées dans cette étude a été accordée par le Comité d'éthique et de bien-être des animaux de l'Université de St Andrews (référence : PS15842). Toutes les données sur les gorilles de montagne et les chimpanzés ont également été collectées avec l'autorisation de l'Uganda Wildlife Authority et du Ugandan National Council for Science and Technology (Conseil national ougandais pour la science et la technologie).

Consentement à la participation Sans objet.

Consentement à la publication Sans objet.

Conflit d'intérêts Les auteurs ne déclarent aucun conflit d'intérêts.

Libre accès Cet article est placé sous licence Creative Commons Attribution 4.0 International License, qui permet l'utilisation, le partage, l'adaptation, la distribution et la reproduction sur n'importe quel support ou dans n'importe quel format, à condition de mentionner les auteurs originaux et la source, de fournir un lien vers la licence Creative Commons et d'indiquer si des changements ont été apportés. Les images ou autres éléments de tiers figurant dans cet article sont inclus dans la licence Creative Commons de l'article, sauf indication contraire dans la ligne de crédit de l'élément. Si le matériel n'est pas inclus dans la licence Creative Commons de l'article et que l'utilisation que vous souhaitez en faire n'est pas autorisée par la loi ou dépasse l'utilisation autorisée, vous devrez obtenir l'autorisation directement auprès du détenteur des droits d'auteur. Pour consulter une copie de cette licence, visitez le site <http://creativecommons.org/licenses/by/4.0/>.

Références

Altmann, J. (1974). L'étude observationnelle du comportement : Sampling meth- ods. *Behaviour*, 49(3-4), 227-266.

- Arnold, K. et Bar-On, D. (2020). Primate pragmatics, expressive behavior, and the evolution of language (La pragmatique des primates, le comportement expressif et l'évolution du langage). *Animal Behavior and Cognition*, 7(2), 117-130.
- Baker, M. (2016). Crise de la reproductibilité. *Nature*, 533(26), 353-366.
- Bain, M., Nagrani, A., Schofield, D., Berdugo, S., Bessa, J., Owen, J., ... Zisserman, A. (2021). Automated audiovisual behavior recognition in wild primates (Reconnaissance automatisée du comportement audiovisuel chez les primates sauvages). *Science Advances*, 7(46), eabi4883.
- Bard, K. A. (1992). Intentional behavior and intentional communication in young free-ranging orangutans. *Child Development*, 63(5), 1186-1197.
- Bard, K. A., Dunbar, S., Maguire-Herring, V., Veira, Y., Hayes, K. G., & McDonald, K. (2014). Gestures and social-emotional communicative development in chimpanzee infants. *American Journal of Primatology*, 76(1), 14-29.
- Bates, E., Camaioni, L. et Volterra, V. (1975). The acquisition of per- formatives prior to speech. *Merrill-Palmer Quarterly of Behavior and Development*, 21(3), 205-226.
- Beery, S., Morris, D. et Yang, S. (2019). Pipeline efficace pour l'examen des images de pièges à caméra. *arXiv preprint arXiv:1907.06772*.
- Boersma, P., & Weenink, D. (2022). Praat : Faire de la phonétique par com- puter [Programme informatique]. Version 6.2.20, récupérée le 24 septembre 2022 sur <http://www.praat.org/>
- Bourjade, M., Cochet, H., Molesti, S. et Guidetti, M. (2020). La diversité con- ceptuelle est-elle un avantage pour l'enquête scientifique ? A case study on the concept of 'gesture' in comparative psychology. *Integrative Psychological and Behavioral Science*, 54(4), 805-832.
- Burghardt, G. M., Bartmess-LeVasseur, J. N., Browning, S. A., Mor- rison, K. E., Stec, C. L., Zachau, C. E., & Freeberg, T. M. (2012). Perspectives- minimiser le biais de l'observateur dans les études comportementales : A review and recommendations. *Ethology*, 118(6), 511-517.
- Cartmill, E. A. et Byrne, R. W. (2007). Orangutans modify their ges- tural signaling according to their audience's comprehension. *Cur- rent Biology*, 17(15), 1345-1348.
- Cartmill, E. A. et Byrne, R. W. (2010). Semantics of primate ges- tures : Intentional meanings of orangutan gestures. *Animal Cog- nition*, 13(6), 793-804.
- Dafreville, M., Hobaiter, C., Guidetti, M., Sillam-Dussès, D., & Bourjade, M. (2021). Sensibilité à l'état attentionnel du partenaire communicant : A developmental study on mother-infant dyads in wild chimpanzees (*Pan troglodytes schweinfurthii*). *American Journal of Primatology*, 83(12), e23339.
- Eaton, T., Hutton, R., Leete, J., Lieb, J., Robeson, A., & Vonk, J. (2018). Du bas vers le haut ! Rejeter les approches top-down centrées sur l'humain en psychologie comparée. *International Journal of Comparative Psychology*, 31.
- ELAN (Version 6.4) [Logiciel]. (2022). Nijmegen : Max Planck Institute for Psycholinguistics, The Language Archive. Extrait de <https://archive.mpi.nl/ta/elan>
- Farrar, B., Krupenye, C., Motes-Rodrigo, A., Tennie, C., Fischer, J., Altschul, D. et Ostojic, L. (2021). Reproduction et reproductibilité dans la recherche sur la cognition des primates
- Favaro, L., Gamba, M., Cresta, E., Fumagalli, E., Bandoli, F., Pilenga, C., ... Reby, D. (2020). Les séquences vocales des manchots obéissent-elles à des lois linguistiques ? *Biology letters*, 16(2), 20190589.
- Ferrer-i-Cancho, R., Bentz, C. et Seguin, C. (2022). Optimal cod- ing and the origins of Zipfian laws. *Journal of Quantitative Linguistics*, 29(2), 165-194.
- Friard, O. et Gamba, M. (2016). BORIS : Un logiciel libre et polyvalent d'enregistrement d'événements pour le codage vidéo/audio et les observations en direct. *Methods in Ecology and Evolution*, 7(11), 1325-1330.
- Fröhlich, M., Kuchenbuch, P., Müller, G., Fruth, B., Furuichi, T., Wittig, R. M., & Pika, S. (2016). Unpeeling the layers of lan- guage : Bonobos and chimpanzees engage in cooperative turn- taking sequences. *Scientific Reports*, 6(1), 1-14.

- Fröhlich, M., van Schaik, C. P., van Noordwijk, M. A. et Knief, U. (2022). Individual variation and plasticity in the infant-directed communication of orang-utan mothers. *Proceedings of the Royal Society B*, 289(1975), 20220200.
- Fröhlich, M. et Hobaiter, C. (2018). Le développement de la communication gestuelle chez les grands singes. *Behavioral Ecology and Sociobiology*, 72(12), 1-14.
- Genty, E. (2019). Combinaisons vocales-gestuelles chez les bébés bonobos : Nouveaux aperçus de la spécificité fonctionnelle du signal. *Animal Cognition*, 22(4), 505-518.
- Genty, E., Breuer, T., Hobaiter, C., & Byrne, R. W. (2009). Communication gestuelle du gorille (*Gorilla gorilla*) : Répertoire, intentionnalité et gestes possibles. *Animal Cognition*, 12(3), 527-546.
- Gilmore, R. O. et Adolph, K. E. (2017). La vidéo peut rendre la science du comportement plus reproductible. *Nature Human Behaviour*, 1(7), 1-2.
- Graham, K. E., Furuichi, T. et Byrne, R. W. (2017). Le répertoire gestuel du bonobo sauvage (*Pan paniscus*) : Un système de communication mutuellement compris. *Animal Cognition*, 20(2), 171-177.
- Graham, K. E., Furuichi, T. et Byrne, R. W. (2020). Context, not sequence order, affects the meaning of bonobo (*Pan paniscus*) gestures. *Gesture*, 19(2-3), 335-364.
- Graham, K. E., Badihi, G., Safryghin, A., Grund, C. et Hobaiter, C. (2022). A socio-ecological perspective on the gestural communication of great ape species, individuals, and social units. *Ethology Ecology & Evolution*, 34(3), 235-259.
- Grund, C., Robbins, M. M., Hobaiter, C. (en préparation) The gestural communication of mountain gorillas.
- Gupta, S. et Sinha, A. (2019). Communication gestuelle des macaques à bonnet sauvages dans le parc national de Bandipur. *Inde du Sud. Behavioural Processes*, 168, 103956.
- Heesen, R., Bangerter, A., Zuberbühler, K., Rossano, F., Iglesias, K., Guéry, J. P., & Genty, E. (2020). Bonobos engage in joint commitment. *Science Advances*, 6(51), eabd1306.
- Heesen, R., Hobaiter, C., Ferrer-i-Cancho, R., & Semple, S. (2019). Lois linguistiques dans la communication gestuelle des chimpanzés. *Proceedings of the Royal Society B*, 286(1896), 20182900.
- Hobaiter, C. et Byrne, R. W. (2011a). Le répertoire gestuel du chimpanzé sauvage. *Animal Cognition*, 14(5), 745-767.
- Hobaiter, C. et Byrne, R. W. (2011b). Serial gesturing by wild chimpanzees : Its nature and function for communication. *Animal Cognition*, 14(6), 827-838.
- Hobaiter, C. et Byrne, R. W. (2017). Qu'est-ce qu'un geste ? Une approche basée sur le sens pour définir les répertoires gestuels. *Neuroscience & Biobehavioral Reviews*, 82, 3-12.
- Hobaiter, C., Badihi, G., Daly, Gabriela, B. D. M., Eleuteri, V., Graham, K. E., Grund, C., Henderson, M., Rodrigues, E. D., Safryghin, A., Soldati, A., & Wiltshire, C. (2021). Base de données vidéo du Dictionnaire des grands singes (1.0.0) [Ensemble de données]. *Zenodo*. <https://doi.org/10.5281/zenodo.5600472>
- Kaufman, A. B. et Rosenthal, R. (2009). Can you believe my eyes ? The importance of interobserver reliability statistics in observations of animal behaviour. *Animal Behaviour*, 78(6), 1487-1491. Keen, S. C., Odom, K. J., Webster, M. S., Kohn, G. M., Wright, T. F. et Araya-Salas, M. (2021). Une approche d'apprentissage automatique pour la classification et la quantification de la diversité acoustique. *Methods in Ecology and Evolution*, 12(7), 1213-1225.
- Kendon, A. (2004). *Gesture : L'action visible comme énonciation*. Cambridge University Press
- Knox, A., Markx, J., How, E., Azis, A., Hobaiter, C., van Veen, F. J., & Morrogh-Bernard, H. (2019). Gesture use in communication between mothers and offspring in wild orang-utans (*Pongo pygmaeus wurmbii*) from the Sabangau Peat-Swamp Forest, Borneo. *International Journal of Primatology*, 40(3), 393-416.
- Leavens, D. A., et Hopkins, W. D. (1998). Intentional communication by chimpanzees : A cross-sectional study of the use of referential gestures. *Developmental Psychology*, 34(5), 813.
- Leavens, D. A., Russell, J. L. et Hopkins, W. D. (2005). Intentionality as measured in the persistence and elaboration of communication by chimpanzees (*Pan troglodytes*). *Child development*, 76(1), 291-306. Liebal, K., Call, J. et Tomasello, M. (2004). Use of gesture sequences in chimpanzees (Utilisation de séquences de gestes chez les chimpanzés). *American Journal of Primatology : Official Journal of the American Society of Primatologists*, 64(4), 377-396.
- Liebal, K., Pika, S. et Tomasello, M. (2006). Gestural communication of orangutans (*Pongo pygmaeus*). *Gesture*, 6(1), 1-38.
- Liebal, K., Slocombe, K. E., & Waller, B. M. (2022). The language void 10 years on : Multimodal primate communication research is still uncommon. *Ethology Ecology & Evolution*, 34(3), 274-287.
- Luef, E. M., et Liebal, K. (2012). Infant-Directed Communication in Lowland Gorillas (*Gorilla gorilla*) : Do Older Animals Scaffold Communicative Competence in Infants ? *American Journal of Primatology*, 74(9), 841-852.
- McCarthy, M. S., Jensvold, M. L. A., & Fouts, D. H. (2013). Use of gesture sequences in captive chimpanzee (*Pan troglodytes*) play. *Animal Cognition*, 16(3), 471-481.
- Mielke, A., Badihi, G., Graham, K. E., Grund, C., Safryghin, A., Hobaiter, C. (en préparation). Many morphs : Establishing great ape gestural repertoires from the bottom-up
- Molesti, S., Meguerditchian, A., & Bourjade, M. (2020). La communication gestuelle chez le babouin olivâtre (*Papio anubis*) : Répertoire et intentionnalité. *Animal Cognition*, 23(1), 19-40.
- Noldus, L. P., Trienes, R. J., Hendriksen, A. H., Jansen, H., & Jansen, R. G. (2000). The Observer Video-Pro : Nouveau logiciel pour la collecte, la gestion et la présentation de données structurées dans le temps à partir de bandes vidéo et de fichiers multimédias numériques. *Behavior Research Methods, Instruments, & Computers*, 32(1), 197-206.
- Collaboration scientifique ouverte. (2015). Estimation de la reproductibilité de la science psychologique. *Science*, 349(6251), aac4716.
- Péter, A. (2011). Solomon Coder (version beta 11.01.22) : une solution simple pour le codage du comportement. *Programme informatique disponible à l'adresse <http://solomoncoder.com>*
- Pika, S., Liebal, K. et Tomasello, M. (2005). Gestural communication in subadult bonobos (*Pan paniscus*) : Repertoire and use. *American Journal of Primatology : Official Journal of the American Society of Primatologists*, 65(1), 39-61.
- Pika, S., Liebal, K. et Tomasello, M. (2003). Gestural communication in young gorillas (*Gorilla gorilla*) : Gestural repertoire, learning, and use. *American Journal of Primatology : Official Journal of the American Society of Primatologists*, 60(3), 95-111.
- Pika, S. et Bugnyar, T. (2011). The use of referential gestures in ravens (*Corvus corax*) in the wild. *Nature Communications*, 2(1), 1-5.
- Pika, S. et Fröhlich, M. (2018). L'acquisition gestuelle chez les grands singes : L'hypothèse de la négociation sociale. *Animal Cognition*, 22(4), 551-565. Pollick, A. S., & De Waal, F. B. (2007). Ape gestures and language evolution. *Proceedings of the National Academy of Sciences*, 104(19), 8184-8189.
- Revelle W (2022). *psych : Procedures for Psychological, Psychometric, and Personality Research*. Northwestern University, Evanston, Illinois. Version 2.2.9 du paquetage R, <https://CRAN.R-project.org/package=psych>
- Roberts, A. I., Vick, S. J., Roberts, S. G. B., Buchanan-Smith, H. M. et Zuberbühler, K. (2012). Un répertoire de gestes manuels basé sur la structure chez les chimpanzés sauvages : Statistical analyses of a graded communication system. *Evolution and Human Behavior*, 33(5), 578-589.
- Rodrigues, E. D., Santos, A. J., Veppo, F., Pereira, J. et Hobaiter, C. (2021). Relier la gestuelle des primates aux racines évolutives du langage : A systematic review. *American Journal of Primatology*, 83(9), e23313.

- Rutz, C. et Webster, M. M. (2021). L'éthologie adopte le cadre STRANGE pour la recherche sur le comportement animal, afin d'améliorer les normes de rapport. *Ethology*, 127(2), 99-101.
- Safryghin, A., Cross, C., Fallon, B., Heesen, R., Ferrer-i-Cancho, R. et Hobaiter, C. (2022). Variable expression of linguistic laws in ape gesture : A case study from chimpanzee sexual solicitation. *Royal Society Open Science*, 9, 220849.
- Schofield, D., Nagrani, A., Zisserman, A., Hayashi, M., Matsuzawa, T., Biro, D., & Carvalho, S. (2019). Reconnaissance de visage de chimpanzé à partir de vidéos dans la nature à l'aide de l'apprentissage profond. *Science Advances*, 5(9), eaaw0736.
- Scott, N. M. et Pika, S. (2012). A call for conformity : Gesture studies in human and non-human primates. *Developments in Primate Gesture Research*, 6, 147.
- Seyfarth, R. et Cheney, D. (2018). Flexibilité pragmatique dans la production vocale des primates. *Current Opinion in Behavioral Sciences*, 21, 56-61.
- Schel, A. M., Bono, A., Aychet, J., Pika, S., & Lemasson, A. (2022). Intentional gestural communication among red-capped mangabeys (*Cercocebus torquatus*). *Animal Cognition*, 1-18.
- Schweinfurth, M. K. et Call, J. (2021). Capuchins (*Sapajus apella*) and their Aversion to Inequity. In *Comparative Cognition* (pp. 173-195). Springer.
- Slocombe, K. E., Waller, B. M. et Liebal, K. (2011). The language void : The need for multimodality in primate communication research. *Animal Behaviour*, 81(5), 919-924.
- Tomasello, M., George, B. L., Kruger, A. C., Jeffrey, M. et Evans, A. (1985). The development of gestural communication in young chimpanzees. *Journal of Human Evolution*, 14(2), 175-186.
- Townsend, S. W., Koski, S. E., Byrne, R. W., Slocombe, K. E., Bickel, B., Boeckle, M., ... Manser, M. B. (2017). Exorciser le fantôme de Grice : une approche empirique pour étudier la communication intentionnelle chez les animaux. *Biological Reviews*, 92(3), 1427-1433.
- Vail, A. L., Manica, A. et Bshary, R. (2013). Referential gestures in fish collaborative hunting (Gestes référentiels dans la chasse collaborative des poissons). *Nature Communications*, 4(1), 1-7.
- Wadewitz, P., Hammerschmidt, K., Battaglia, D., Witt, A., Wolf, F., & Fischer, J. (2015). Caractérisation des répertoires vocaux - approches de classification dures et douces. *PloS One*, 10(4), e0125785.
- Webster, M. M. et Rutz, C. (2020). How STRANGE are your study animals ? *Nature*, 337-340.
- Wiltshire, C., Lewis-Cheetham, J., Komedová, V., Matsuzawa, T., Gramham, K. E., & Hobaiter, C. (en cours de révision). DeepWild : Application of the pose estimation tool DeepLabCut for behaviour tracking in wild chimpanzees and bonobos. *Journal of Animal Ecology*
- Zipf, G. K. (1949). Human behavior and the principle of least effort : An introduction to human ecology. Cambridge

Note de l'éditeur Springer Nature reste neutre en ce qui concerne les revendications juridictionnelles dans les cartes publiées et les affiliations institutionnelles.