

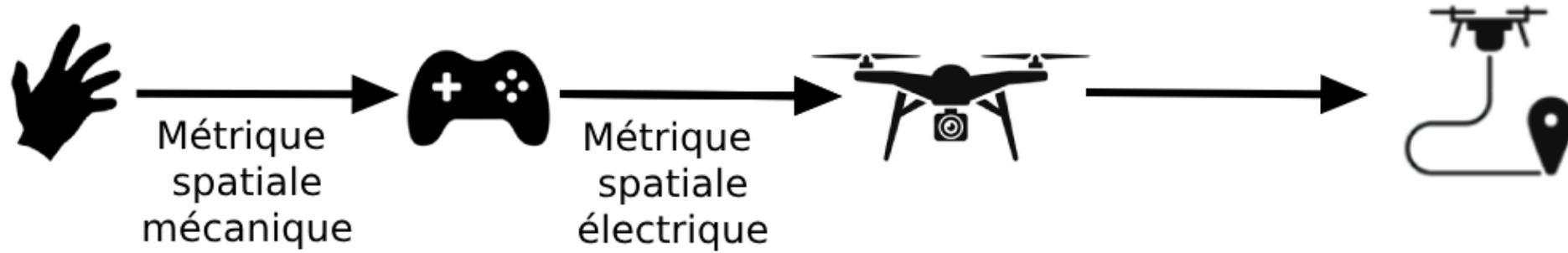
# Commandes vocales spontanées et déplacement guidé vers une position 3D : première étape pour la navigation de drone

HENRY Allan

18 septembre 2024

# Contexte

## Méthode classique



**Problème : L'utilisateur est obligé d'apprendre les commandes avant de pouvoir piloter le drone**

**Objectif : N'importe qui puisse utiliser un drone sans avoir de formation préalable**

## Commande vocale de drone

- Liste de commandes prédéfinies (6 et 200 mots) [1][4][5]
  - S'intéresse aux mots prononcés, mais pas à la manière dont ils sont prononcés
  - Création d'une grammaire à apprendre avant l'utilisation [2]
    - Ex. [uncertain velocity] [verb] [uncertain distance] [determinant] [object] [uncertain velocity]
- Deux types de commandes [3]
  - Bas niveau : Va à gauche
  - Haut niveau : Va à l'arbre
- Utilisation de Large Language Model (LLM) [6]
  - Permet d'interpréter les instructions pour qu'elles correspondent à une commande



[1] Jeong-Sik Park et Na Geng. "In-Vehicle Speech Recognition for Voice-Driven UAV Control in a Collaborative Environment of MAV and UAV"

[2] Samantha Rajapaksha et al. "Responsive Drone Autopilot System for Uncertain Natural Language Commands"

[3] Geraint Jones et al. "Towards a situated, multimodal interface for multiple UAV control"

[4] Choutri Kheireddine, Mohand Lagha, Souham Meshoul, Mohamed Batouche, Yasmine Kacel and Nihad Mebarkia. "A Multi-Lingual SpeechRecognition-Based Framework to Human-Drone Interaction."

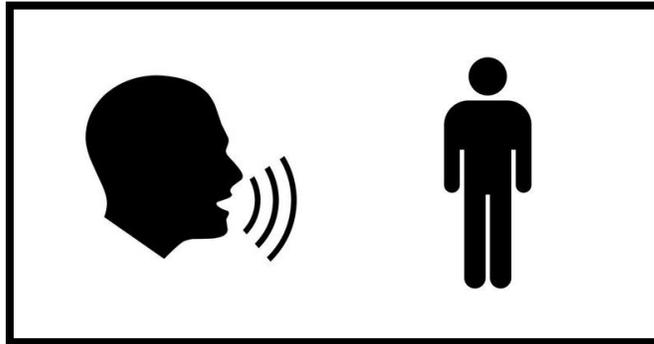
[5] Poncela Alberto and Leticia Gallardo-Estrella. "Command-Based VoiceTeleoperation of a Mobile Robot via a Human-Robot Interface."

[6] Nwankwo, Linus and Elmar Rueckert. "The Conversation Is the Command: Interacting with Real-World Autonomous Robots Through Natural Language."

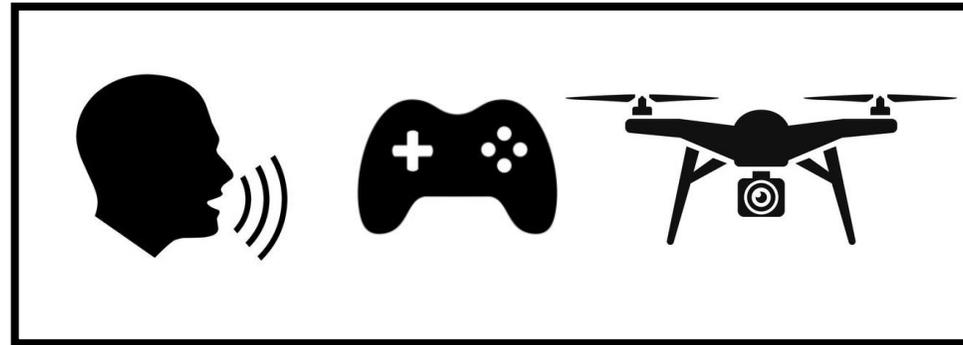
## Voir l'humain comme un concepteur et pas uniquement comme un usager

- Ne donner aucune consigne à l'humain
- Ne pas imaginer les catégories de commandes mais voir celles qui émergent
- Comment adapter le drone à l'humain

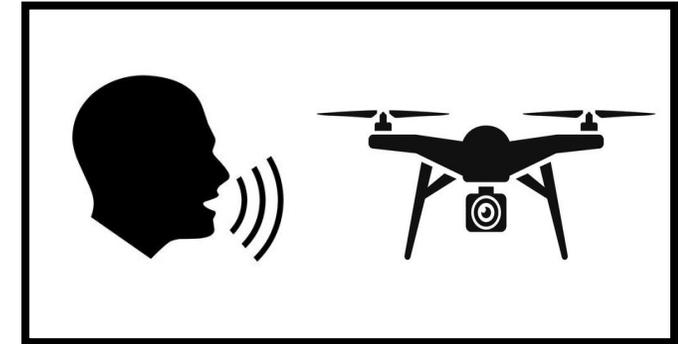
## 3 expériences



**Guide - Humain**

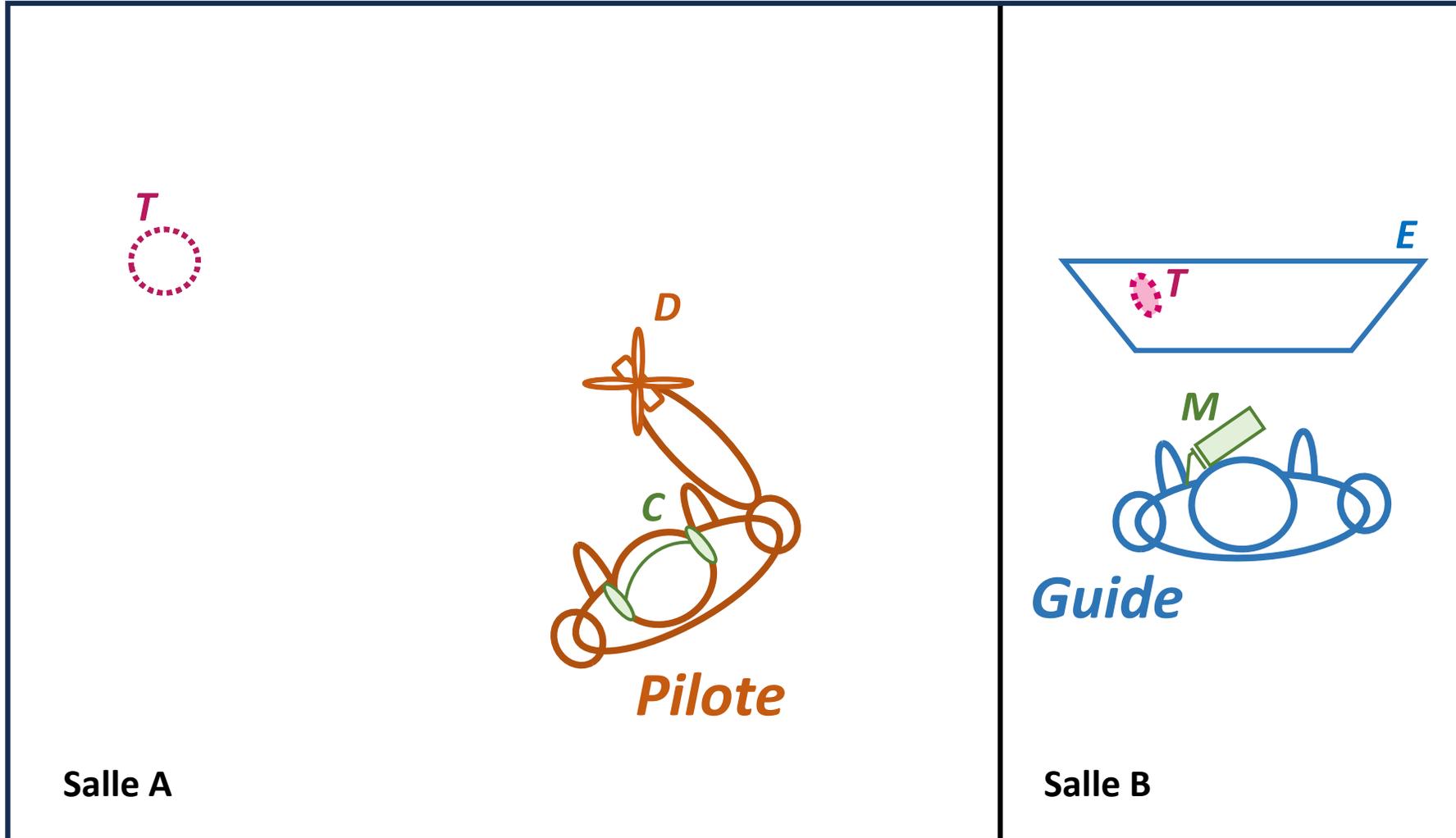


**Guide - Pilote de drone**



**Guide - Drone**

# Expérience Guide - Humain

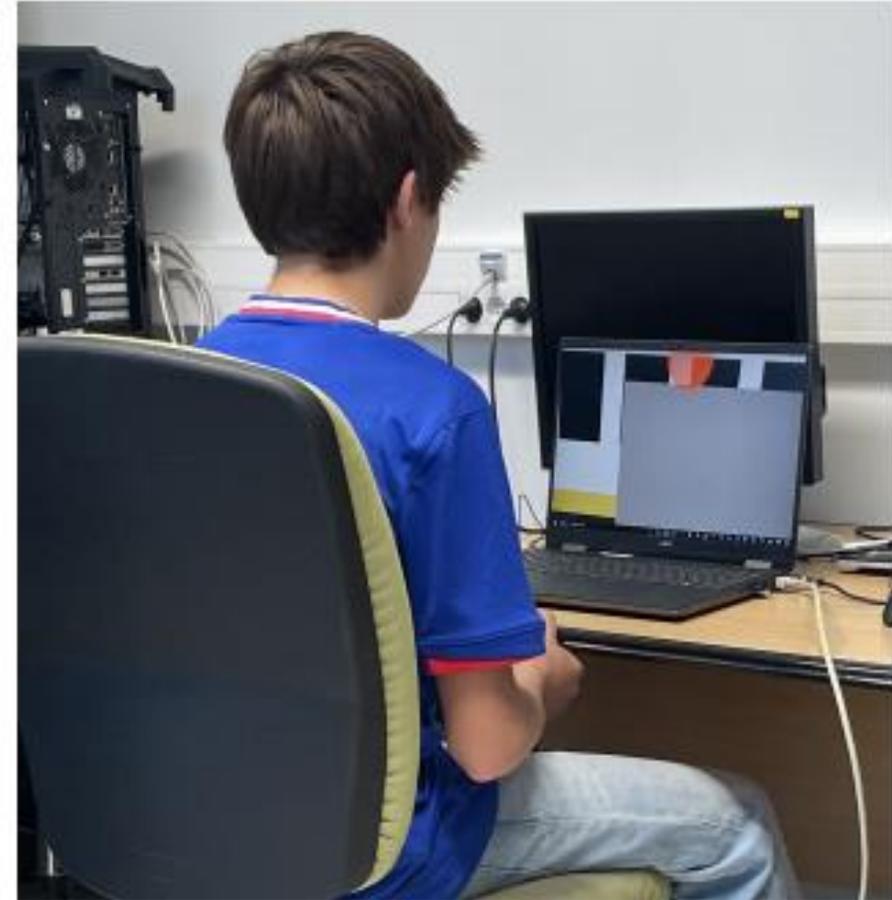


- D** : Drone (DJI Tello)
- T** : Cible
- C** : Casque
- M** : Microphone
- E** : Ecran

# Expérience Guide - Humain

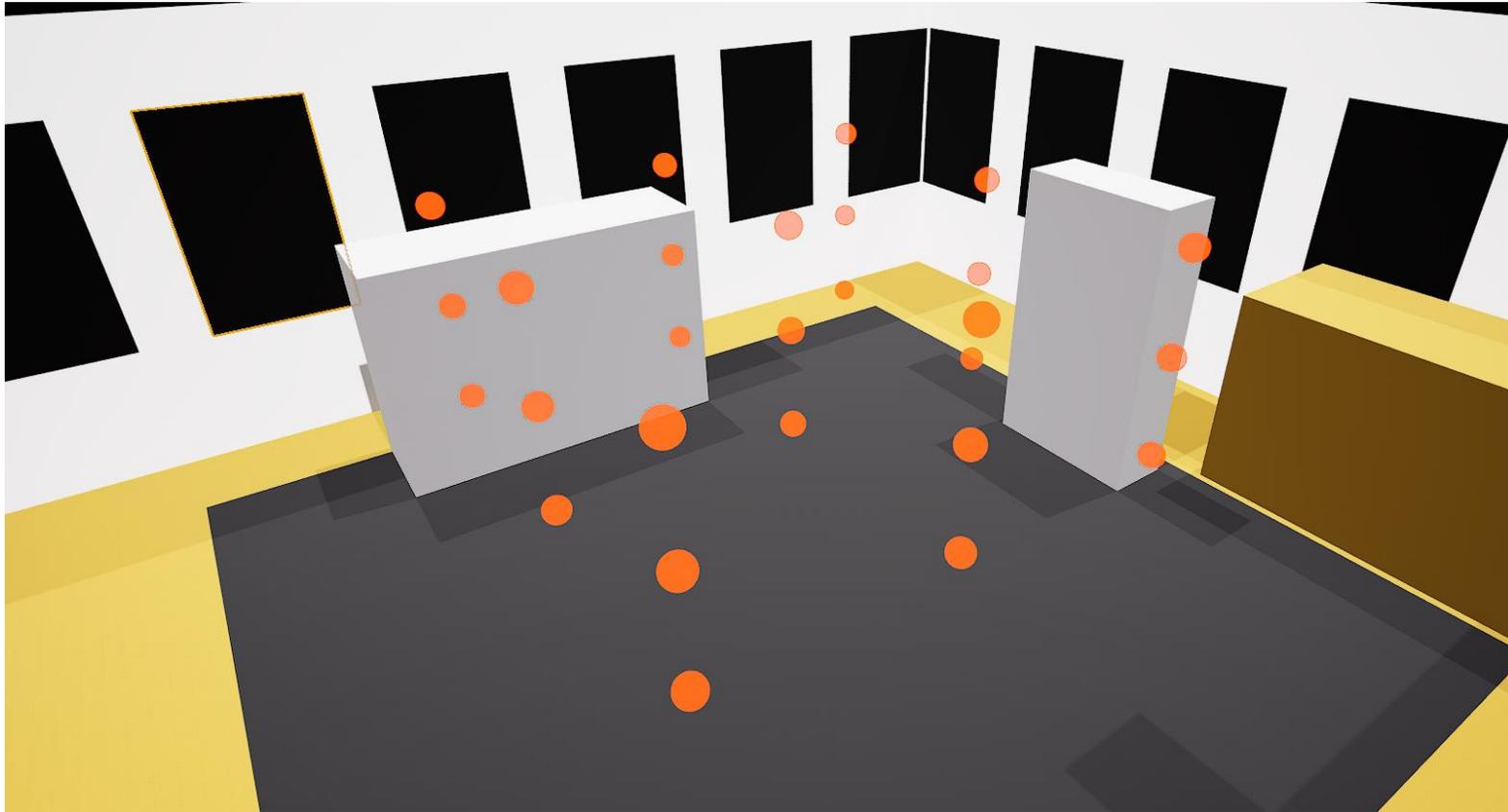


Salle A



Salle B

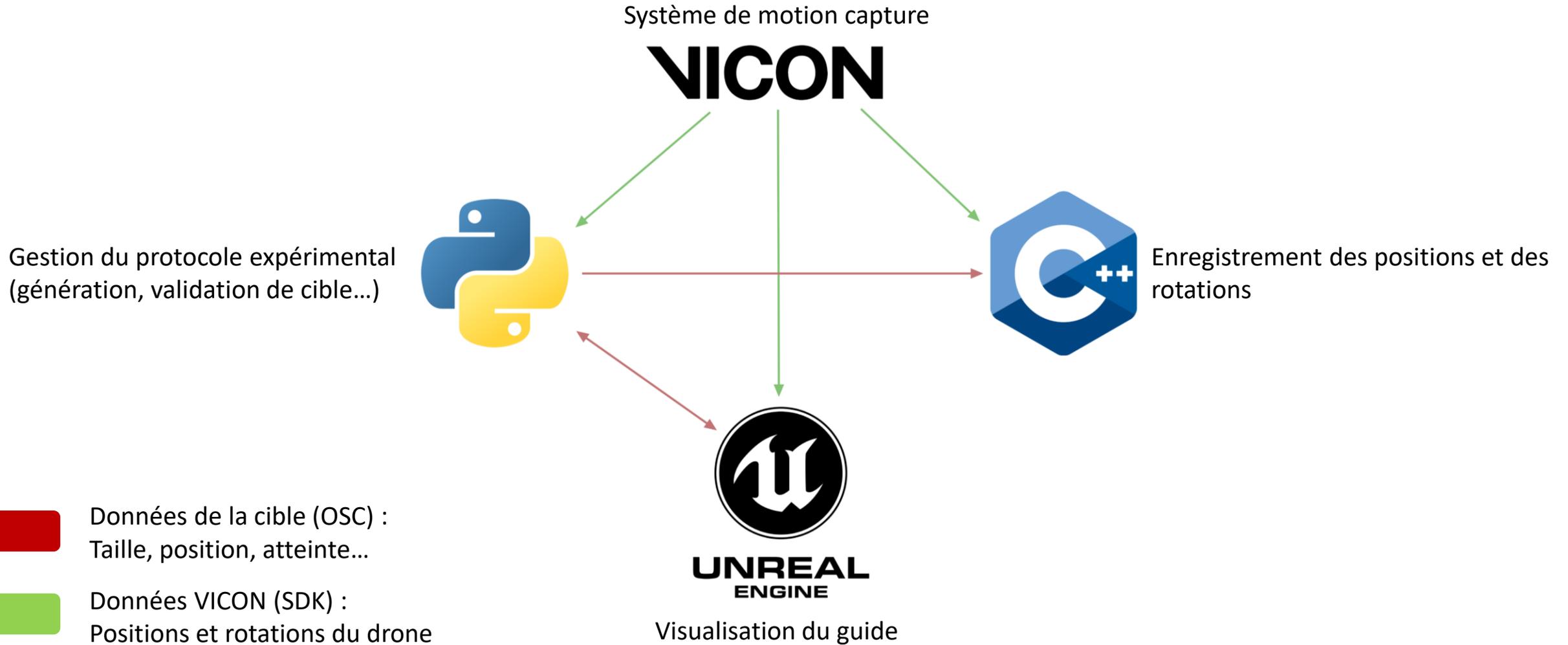
# Génération de cible



## Les cibles :

- Sphères de 30 cm de diamètre
- Sur une grille de 3x3x3
- Un espacement de 135 cm en X et en Y.
- Trois hauteurs différentes : 70 cm, 110 cm et 150 cm
- 6 cibles à atteindre
- Temps maximal de 120s par cible

# Architecture logicielle



# Participants

- 18 participants (6 hommes, 10 femmes et 2 non-binaires),
- Âgés de 15 à 25 ans ( $M = 18.92$ ,  $SD = 4.14$ )
- Ayant peu d'expérience dans les activités de guidages (animal de compagnie, activité collective, activité coachée)

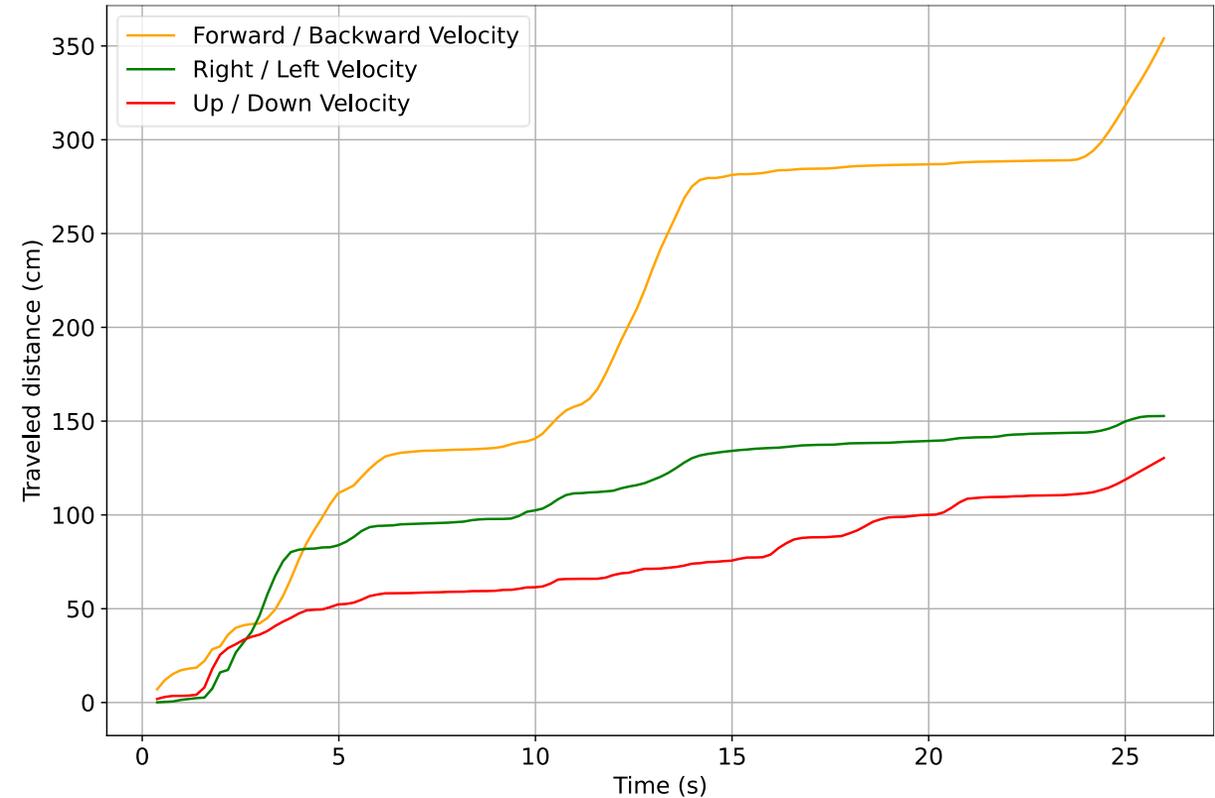
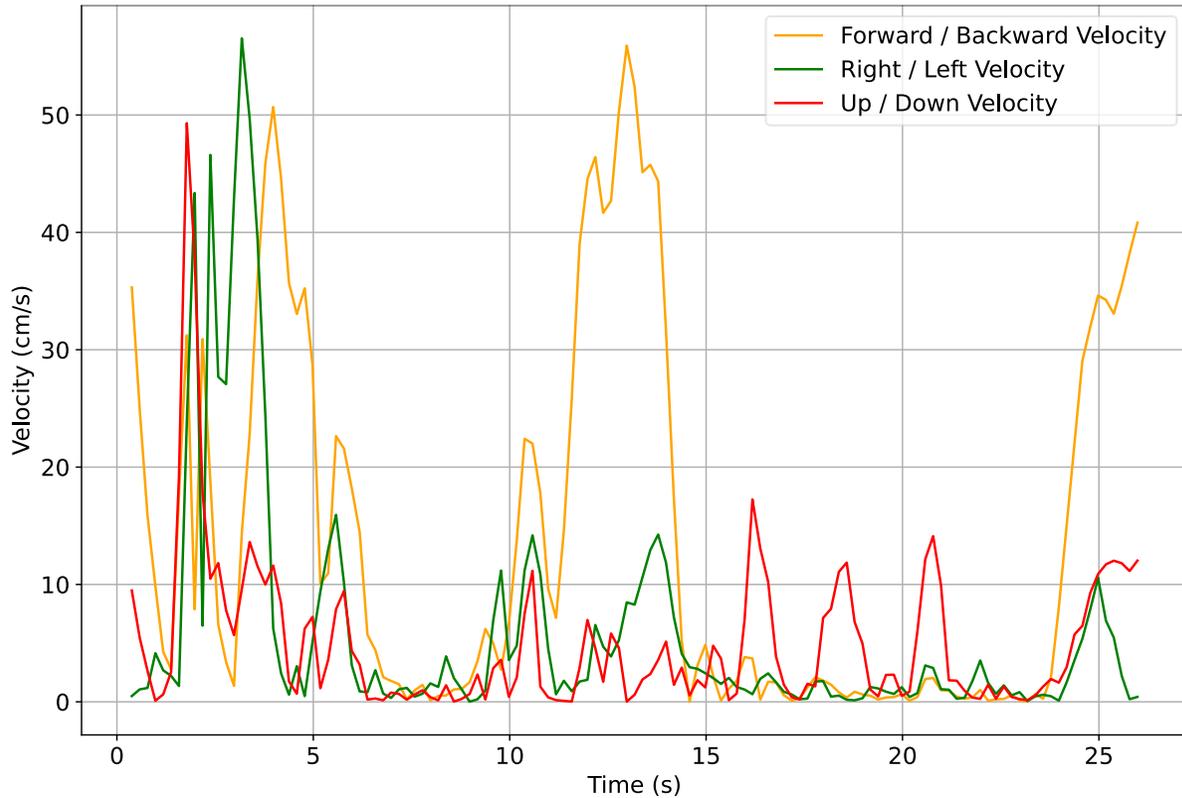
Bloc	Guide	Pilote	Repos
1	A	B	C
2	B	C	A
3	C	A	B
4	A	C	B
5	B	A	C
6	C	B	A

# Résultats

- 18 participants = 216 cibles
  - 12 cibles avec des données erronées (problème de périphérique)
  - 11 cibles non atteinte : toutes lors des trois premiers blocs
  - 193 cibles atteintes

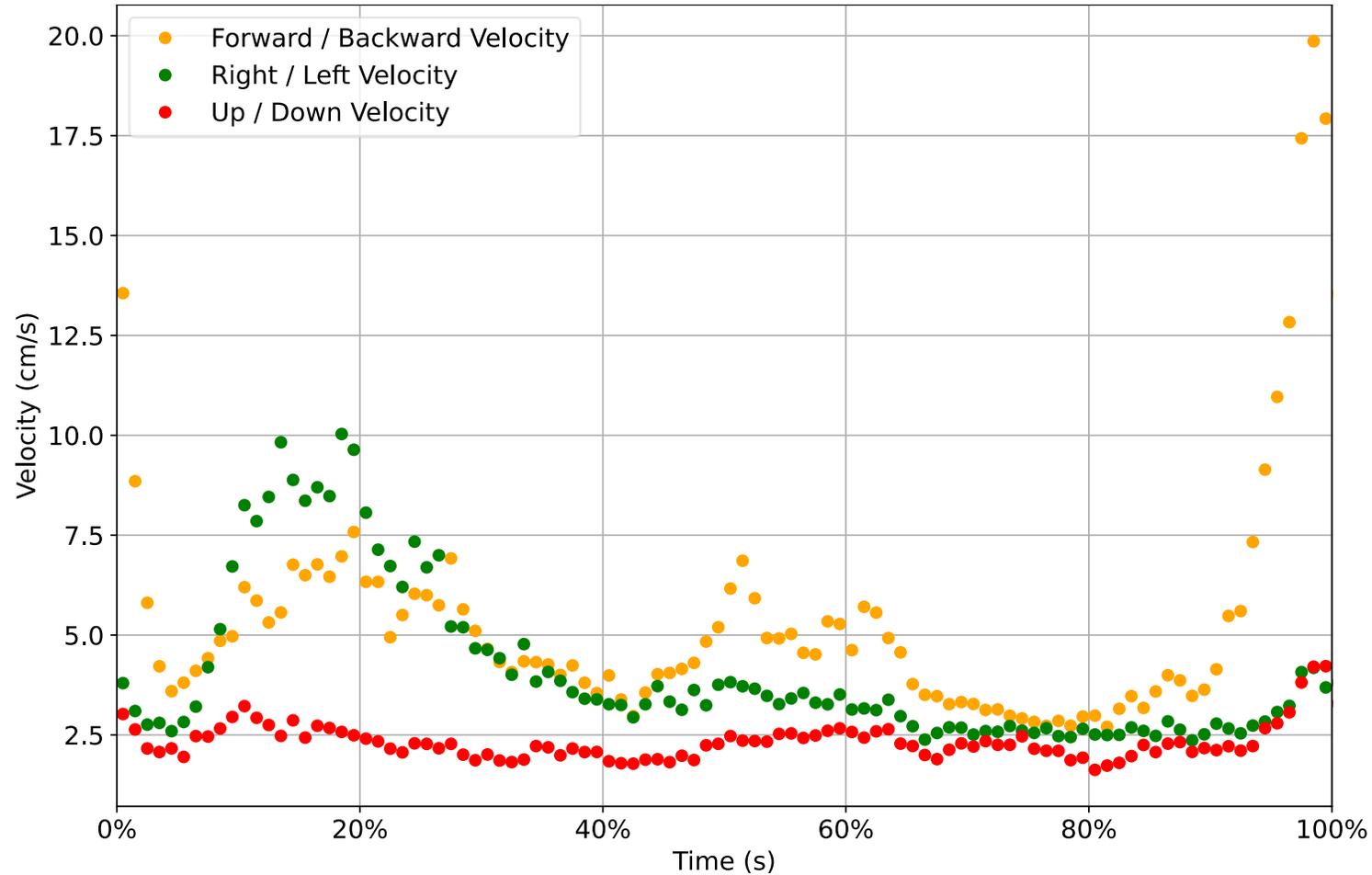
	Min	Max	Moyenne	Médiane	Ecart-type
Durée (s)	13,8	120	35,75	42,92	25,45
Distance parcourue (cm)	191	3183	574	704	435

## Vitesses et distance parcourue d'un essai selon les axes égocentriques du drone



# Résultats

## Vitesses médianes de tous les essais en fonction du temps normalisé en pourcentage



## La trajectoire peut être divisée en 3 phases

<b>1. Orientation</b>	L'humain se tourne jusqu'à ce que le guide voie la cible
<b>2. Approche</b>	L'humain se rapproche de la cible
<b>3. Ajustement</b>	La cible est à portée de l'humain

# Résultats

## Les 3 phases se distinguent par le lexique utilisé Nombre d'occurrences sur 72 essais

	1. Orientation		2. Approche		3. Ajustement	
	Tourner	58	Stop	104	Avancer	120
	Stop	34	Avancer	94	Stop	115
	Sur	28	Un peu	67	Un peu	85
	Même	27	Encore	57	A gauche	69
	Là	25	A gauche	43	Encore	68
	Cible	18	Tout droit	32	C'est bon	64
	C'est bon	13	Tourner	29	Là	64
<b>Nombre moyen de mots par phase</b>	11		21		27	
<b>Durée médiane entre les instructions (s)</b>	0,22		0,26		0,18	

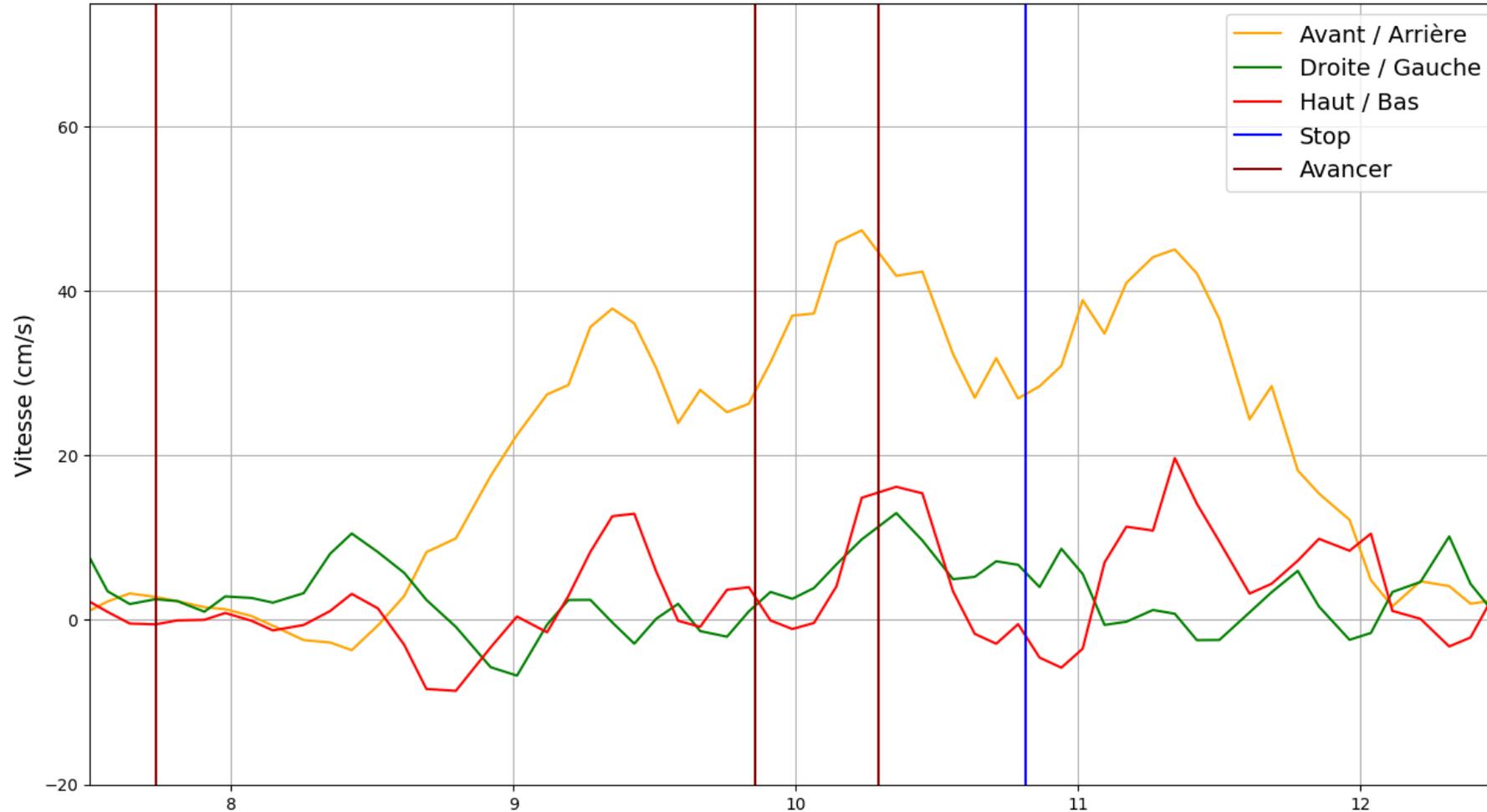
# Pistes de recherches

## Travaux en cours

- Classification des données (commandes et trajectoires)
- Etude des répétitions et des aspects prosodiques

# Résultats préliminaires

## Vitesses d'un essai selon les axes égocentriques du drone



# Pistes de recherches

## Travaux en cours

- Classification des données (commandes et trajectoires)
- Etude des répétitions et des aspects prosodiques

## Perspectives

- Expérience guide / pilote de drone
  - Dynamique des trajectoires => commandes différentes ?
- Expérience guide / drone
  - Utilisation des résultats précédents

Merci de votre attention !