Full Action Scene Graph - FASG

CAMERA WEARER → *agent* → **HOLD**

HOLD → *direct object* → **CUP** (attr: large)

HOLD → *aux verb* → **SIT**

CUP → *of* → **COFFEE** (attr: black)

SIT → *aux direct object* → **TABLE**
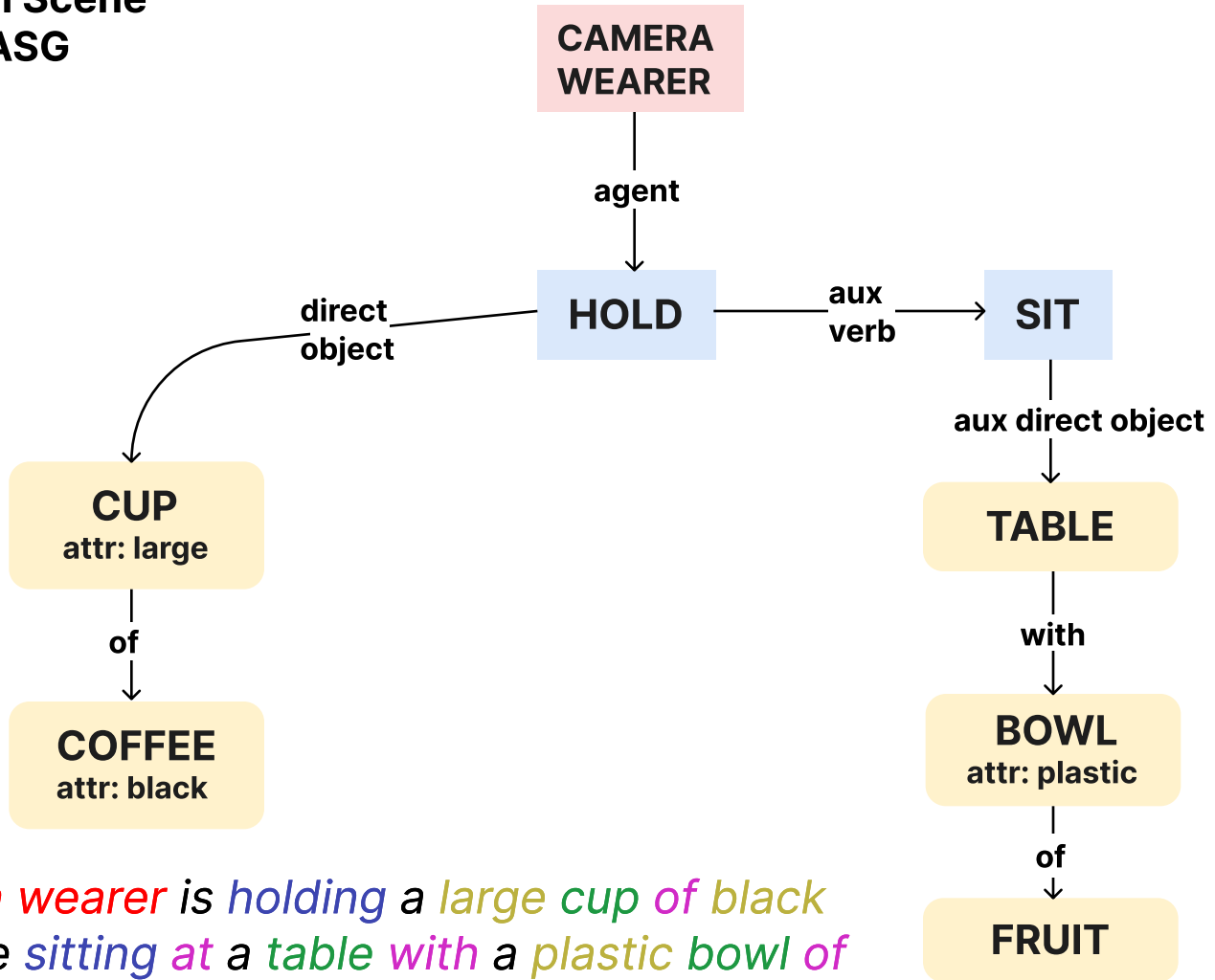
TABLE → *with* → **BOWL** (attr: plastic)

BOWL → *of* → **FRUIT**

The camera wearer is holding a large cup of black coffee while sitting at a table with a plastic bowl of fruit.