Interlocutors converge on names to refer to entities. For example, a speaker might refer to a novel looking object as the jellyfish and, once identified, the listener will too. The hypothesized mechanism behind such referential precedents is a subject of debate. The common ground view claims that listeners register the object as well as the identity of the speaker who coined the label. The linguistic view claims that, once established, precedents are treated by listeners like any other linguistic unit, i.e. without needing to keep track of the speaker. To test predictions from each account, we used visual-world eyetracking, which allows observations in real time, during a standard referential communication task. Participants had to select objects based on instructions from two speakers. In the critical condition, listeners sought an object with a negative reference such as not the jellyfish. We aimed to determine the extent to which listeners rely on the linguistic input, common ground or both. We found that initial interpretations were based on linguistic processing only and that common ground considerations do emerge but only after 1000 ms. Our findings support the idea that—at least temporally—linguistic processing can be isolated from common ground.

1. Introduction

When a mother glances at her son, says ‘Look’ and points ahead, she expects her son to understand that her intention was to