Please quote the published version
Grodniewicz, J. P. (2021). The process of linguistic understanding. *Synthese*, *198*(12), 11463–11481.
https://doi.org/10.1007/s11229-020-02807-9

The process of linguistic understanding

J.P. Grodniewicz
(j.grodniewicz@gmail.com)

Abstract. The majority of our linguistic exchanges, such as everyday conversations, are divided into turns; one party usually talks at a time, with only relatively rare occurrences of brief overlaps in which there are two (or more) simultaneous speakers. Moreover, conversational turn-taking tends to be very fast. We typically start producing our responses before the previous turn has finished, i.e., before we are confronted with the full content of our interlocutor's utterance. This raises interesting questions about the nature of linguistic understanding. Philosophical theories typically focus on linguistic understanding characterized either as an ability to grasp the contents of utterances in a given language or as outputs of this ability—mental states of one type or another. In this paper, I supplement these theories by developing an account of the process of understanding. I argue that it enables us to capture the dynamic and temporal aspect of understanding and reconcile philosophical investigations with empirical research on language comprehension.

Keywords: linguistic understanding; turn-taking; what is said; language processing; knowledge of language

1. Introduction

According to conversation analysis studies, the majority of our linguistic exchanges, such as everyday conversations, are divided into turns (Sacks et al., 1974). That is to say that one party usually talks at a time, with only relatively rare occurrences of brief overlaps in which there are two (or more) simultaneous speakers. This observation would not be striking if it were not juxtaposed with another empirical result provided by psycholinguistics: the gaps between subsequent turns are very short, only about 0.2 s (Stivers et al., 2009). Given that language production requires a minimum of around 1 s for an unprimed single-word utterance to around 1.5 s for a simple clause, we have to conclude that subsequent turns are planned, and that the production process begins before the end of the previous turn (Levinson and Torreira, 2015; Garrod and Pickering, 2015; Levinson, 2016). Consequently, speakers typically start responding not to the whole content of an utterance they have heard, but to their prediction of the full content of the utterance on the basis of only that part of the content they have actually heard.[1] To account for these data, we need a

[1] Of course, to say that 0.2 s is the length of a typical gap in conversation is a vast generalization. A lot depends on such factors as the type of the upcoming turn. For example, Kendrick (2015) observes that turns involving repair appear approximately 0.7 s after the preceding ones. However, even 0.7 s is still a short enough gap to support the argument presented in what follows.

theory of linguistic understanding more versatile and of a broader scope than those already available in the philosophical debate.[2]

Philosophical theories typically focus on states of understanding, i.e., "outputs of exercises of one's ability" to understand (Longworth, 2018). They look for an answer to a very specific question: What is the hearer's attitude towards the content or *what is said* through an utterance which that hearer understands? Most influential theories characterize understanding either as a kind of: knowledge (e.g., Evans, 1982; Dummett, 1993; Davies, 1989; Heck, 1995; Campbell, 1982; Higginbotham, 1992), belief (e.g., Millikan 1984, 2004), perception-like state of direct awareness of the content (and force) of the utterance (Hunter 1998, Fricker 2003), or a state through which the hearer entertains the content of the utterance (Longworth, 2018).

Research on turn-taking shows that we start producing our responses before the turn in progress at the time has finished, i.e., before we are confronted with the full content of our interlocutor's utterance. Certainly, it is quite natural to assume that a hearer is in a state of understanding an utterance as soon as the last word of that utterance has been heard. However, short turn-taking gaps simply do not provide enough time for a hearer to produce a response only when already in a state representing the content of the previous turn. Thus, apparently, in the vast majority of communicative situations we produce our conversational turn without being in one of the states typically characterized by philosophers as the state of understanding (I come back to this matter in Section 4). In result, to account for the role it plays in fast communicative exchanges, we have to extend our characterization of linguistic understanding beyond the most popular state-sense. The authors enumerated above do not claim that being in a state of understanding is *all* there is to linguistic understanding. However, none of them provide an elaborated account of the dynamic aspect of linguistic understanding. This paper is meant as an attempt to provide such an account.[3]

It is important to keep in mind that the current paper is not meant as a critique of the available theories of linguistic understanding listed above. The states of understanding that they focus on are independently interesting. We want such states to play various roles in our cognitive enterprises. For example, states of understanding figure as partial reasons for action (e.g., I have run away because I knew that what you have said was that there is a tiger behind the rock), and belief formation (I believe that there is a tiger behind the rock because this is what you have told me). Some theory of states of understanding has to be a part of a more versatile theory of linguistic understanding I am about to offer.[4]

My plan is the following. In Section 2, I present three senses of linguistic understanding. In Section 3, I discuss the relation between these senses and develop an account of the process of understanding. In Section 4, I respond to a possible objection according to which we can account for the dynamic aspect of linguistic understanding without characterizing linguistic understanding

[2] It is important to keep in mind that linguistic understanding that is the subject of this paper should not be identified with communicative success. Arguably, communicative success may, in specific situation, be achieved despite the lack of linguistic understanding (or in the case of only partial understanding). Thus, linguistic understanding is not identical with the communicative success, although might (and typically does) contribute to it.
[3] In this essay, I focus on the understanding of speech (spoken linguistic input). I think that it will be relatively easy to extend my proposed account to the written word, although I do not undertake this task here.
[4] I outline my view on states of understanding in Author XXX. Moreover, in Section 4, I discuss the hypothesis that the process of understanding is constituted out of states of (partial) understanding.

as a process. Finally, in Section 5, I outline additional benefits of my account for the philosophical debate about linguistic understanding.

2. Three senses of "linguistic understanding"

There are at least three senses in which we pre-theoretically grasp the idea of linguistic understanding (Hunter, 1998; Longworth, 2010). I have already mentioned the first: we often think of understanding as a mental state in which a subject represents the content of a heard utterance. In the second sense, we consider understanding to be a disposition or stable ability to decode sentences of a certain language (I will be sometimes referring to it as *disposition-understanding* of a given language). This is what we mean when we talk of understanding (or simply "speaking") a given language, such as English, Kashmiri, or Northern Sotho.

It is worth noting that we expect states of understanding (the first sense of understanding given above) to be grounded in a disposition (this second sense of understanding). Fricker (2003) highlights the fact that to have an understanding of a particular utterance of a sentence, S, is to enjoy a correct representation of the content of the utterance "as a result of one's dispositional understanding S—that is, through the exercise of a stably possessed, internally constituted reliable capacity" (p. 346). It is not enough to simply grasp the content of an utterance. As Fricker shows through her example, I might effectively grasp the contents of Russian utterances if there is a reliable translator by my side who translates from Russian to English for me. Does that mean that I understand Russian? In a sense, I do; in a sense, I do not. If, asking about understanding, we ask whether I can communicate with my Russian monolingual friend, it may seem that I understand Russian (thanks to my Russian translator!). However, we might be concerned with a more profound matter: whether I myself as a speaker with such and such biological cognitive setting, such-and-such experience and language acquisition history understand Russian. In this latter case the answer would be "no". It is obvious that what is of primary importance for our research into the nature of linguistic understanding is the latter question, and the sense in which I do not understand Russian. At least when trying to refine the basic sense of linguistic understanding (which could, later on, be used to set up standards for possible derivative senses such as "I don't know any Russian but I could understand him because I had Google Translate on my phone."), we need to focus on a disposition to understand a language which is common to almost all human beings and which is internalized in a very specific way. It is grounded in our language faculty (e.g., Hauser et al., 2002; Jackendoff and Pinker, 2005) consisting, among others, of an appropriate biological setting of a language-ready brain (e.g., Boeckx and Benítez-Burraco, 2014; Friederici, 2012), some form of internalization of the grammar of a given language (e.g., Pinker and Bloom, 1990), a mental lexicon (e.g., Jackendoff, 2002; Patterson et al., 2007), etc. Due to the lack of space, I will leave a discussion of the exact nature of that disposition for another occasion. However, I agree with Fricker that possession of such disposition is necessary for being in states of understanding of particular utterances. States of understanding of utterances of sentences in a given language result, at least in part, from exercising the disposition to understand this language.

The third way in which we can think of understanding is to think of it as an "event" (Hunter, 1998, 561), "process" (Fricker, 2003, 330), "episode" or "achievement" (Longworth, 2010, 4) of one's

coming to the state of understanding of a given utterance.[5] Notably, there is relatively little discussion about this sense of understanding, either in the philosophy of language or the philosophy of mind. Someone might even assume, that the nature of processes, events, or episodes of understanding lies outside the area of philosophical considerations; it is the task of psycholinguistics to study time-spans and the mechanisms of language comprehension, and that of neurolinguistics to study the neural pathways recruited for the processes, etc. However, as I argue in the next section, focusing on the dynamic and temporal character of linguistic understanding opens a rich and interesting avenue of philosophical research.

3. Linguistic understanding as a cognitive process

3.1 The metaphysics of processes

How do the three senses of linguistic understanding mentioned above hang together? Are they all pointing at entirely distinct phenomena or rather different aspects of the same complex phenomenon? Finally, is it possible that careful examination of any one of these senses will help us to better understand the remaining two?

I think that the three senses of linguistic understanding indicate different aspects of the same phenomenon, and as such, they are strongly interdependent. I will now propose a stipulative definition of linguistic understanding, which outlines the relations between them. Given that my primary interest in the present paper is the dynamic and temporal aspect of linguistic understanding, my definition puts the process sense of linguistic understanding to the foreground. Crucially, the definition is, at least prima facie, compatible with a variety of views about the nature of state-understanding and disposition-understanding available in the debate.

> **(LU)** Linguistic understanding is a cognitive process that typically takes place on hearing an utterance in a language one disposition-understands and which typically results in a set of mental states *via* which the hearer represents the content and force of the utterance or what the hearer identifies as the content and force thereof.

LU assumes very little about the nature of the process of understanding. In principle, it is possible that, e.g., Cyborgs or Martians were capable of performing *the same* process, even if they were equipped with different cognitive and perceptual mechanisms than these possessed by humans. Independent of a specific implementation, a process that is based on a disposition to understand a given language and produces mental states representing contents of utterances of sentences in this language, qualifies as a process of understanding.

Moreover, LU specifies what triggers the process of linguistic understanding. Just as the process of watching a particular material object requires a particular material object that is being watched (Crowther, 2009), the process of linguistic understanding requires some linguistic input that is being understood. However, restriction to *typical* cases is important here. Not always when a subject possessing a disposition-understanding of a given language encounters utterances in this language, the process of understanding is being triggered. If the hearer is cognitively impaired or if their cognitive system has been tinkered with or, finally, if their attention resources are

[5] I will discuss the differences between these notions in the next section.

insufficient it is possible that they would encounter an utterance but the process would not be triggered or the process would be triggered but the state would not be produced (more on this below). As is the case with other types of processes dependent on the working of our cognitive mechanisms, we should restrict our predictions about linguistic processing to broadly construed *normal conditions*.

But is it appropriate to think about linguistic understanding as a process? To answer this question, I will draw from a rich tradition of philosophers who investigated the nature of temporal entities, often in the context of the discussion about grammatical aspect (cf. Vendler, 1957; Verkuyl, 1972; Mourelatos, 1978; O'Shaughnessy, 2000; Rothstein, 2004; Crowther, 2009, 2011, 2018; Steward 1997, 2012, 2018; Galton, 2018; Soteriou, 2013, 2018; Stout, 2018). Different authors use different classifications but in the present paper, I will distinguish four basic categories: states, achievements, accomplishments, and activities. My goal in this section is to argue that there are senses of "linguistic understanding" belonging to each of these categories. Is it even possible? It is. As indicated by Vendler:

> [it is not the case that] a verb exhibiting a use fairly covered by one schema cannot have divergent uses, which in turn may be described in terms of other schemata. As a matter of fact, precisely those verbs that call for two or more time schemata will provide the most interesting instances of conceptual divergence in this respect — an ambiguity which, if undetected, might lead to confusion. Thus my intention is not to give rules about how to use certain terms but to suggest a way of describing the use of those terms. (Vendler, 1957, 143-4)

As I have already mentioned, there is an obvious sense in which "linguistic understanding" denotes a *state*. Depending on which theory we favor, we can characterize the state of understanding of an utterance either as knowledge that what was said was so and so, a belief that what was said was so and so, a state of entertaining the content of this utterance, etc. Crucially, "states such as knowledge and belief do not have temporal duration in the sense of having temporal parts or successive temporal phases over which they unfold." (Crowther, 2011, 5). The same state of my understanding of your utterance that *p* may obtain now and in an hour, and it does not make sense to say that in an hour I am in a different phase, or at a different stage of this state.

Importantly, disposition-understanding which we pick out when we say "Anne understands Italian" is also a state, although different than a state of understanding of a particular utterance. As I said above, disposition-understanding of a given language is a state of a subject's cognitive system, which enables the system to respond to linguistic input, e.g., particular utterances, by producing states of understanding of these utterances.

The next category enumerated above is this of achievements. Achievements are typically characterized as instantaneous changes of state. Some clear examples are *noticing*, *recognizing*, *finding*, and *dying*. As observed by Longworth (2010), there is an intuitive sense in which *linguistic understanding* denotes an achievement. In this sense, understanding is an episode of: "one's coming to state-understand, or understanding, what someone has said" (Longworth, 2010, 4); it is the temporal boundary between whatever led to the production of state-understanding and the first moment in which the state-understanding obtains.

Linguistic understanding has been previously characterized as a state and as an achievement, so these two senses are fairly uncontroversial. But I have promised to argue that there are senses of linguistic understanding corresponding to all four categories enumerated above, that is also to activities and accomplishments. In what follows, I will use the term "process" as a collective name for activities and accomplishments.[6] Therefore, my goal now is to argue that we can think about linguistic understanding as a process.

Activities and accomplishments are both progressive, i.e., they describe temporal entities unfolding over time. Names of particular activities and accomplishments can be appropriately used as answers to questions such as "What are you doing?" or "What is going on?" (Steward, 2018, 111). The core difference between activities and accomplishments concerns telicity.[7] Accomplishments such as *walking to the shops* or *writing a letter* are oriented towards specific goals or points "at which they must terminate" (Crowther, 2011, 6), while activities, such as *walking* or *writing*, are not. If I have walked for ten minutes and after that kept walking for another five minutes, there was an activity of walking that took fifteen minutes. However, if I have walked to the shops in ten minutes, it means that I have reached the shops after ten minutes. After that, I can walk to the shops *again*, but I cannot walk to the shops *some more*.[8]

Understanding an utterance is a candidate for an accomplishment, because it has specific "completion condition" (Crowther, 2011), namely, the achievement of the state of understanding. The claim I want to defend is that *understanding an utterance* unfolds over time because *hearing an utterance* unfolds over time and the temporal structure of understanding an utterance overlaps with the temporal structure of hearing this utterance.[9] In Section 3.2, I present psycholinguistic data about incrementality and predictivity of language understanding, which support this claim. Here, my focus is on its philosophical motivation.

Let us consider the following, particularly interesting example of an accomplishment provided by Matthew Soteriou (2018):

> It takes around 51 minutes and 14 seconds to play, from start to finish, the 1981 recording of Glenn Gould's performance of the Goldberg Variations. It takes at least the same amount of time to hear it. Once you've listened to the whole recording, you will have heard the final recording of Gould's performance of that work… Although we're unable to hear a performance of the complete work all at once, so to speak, this doesn't mean that we are incapable of hearing a performance of the whole work. It just means that this is an accomplishment that takes time. (Soteriou, 2018, 82)

The reason why it takes 51 minutes and 14 seconds to hear the 1981 recording of Goldberg Variations is that playing this piece takes the same amount of time. At each moment of playing, there is simultaneous hearing. But, let's go one step further and imagine that the hearer in Soteriou's example is a music critic whose goal is to analyze the 1981 recording. Most likely, the

[6] Other authors pick different terminological conventions. For example, Crowther (2011) uses the term "process" for what I, following e.g., Vendler (1957) and Rothstein (2004), call "activity."
[7] Cf. Rothstein (2004).
[8] For an exhaustive and illuminating discussion of the difference between activities and accomplishments see (Crowther, 2011).
[9] That does not mean that they have *the same* temporal structure. More on this topic in Section 4.

critic does not wait with their analysis till the *playing* is finished and the *hearing* accomplished. They analyze Goldberg Variations as they hear them being played. So, it takes (at least) 51 minutes and 14 seconds to analyze the 1981 recording of Goldberg Variation.

I submit that, in this respect, *understanding an utterance* is analogous to *analyzing the 1981 recording of Goldberg Variations*.[10] If you utter the sentence "It would be really nice to visit Vienna again, like in the old times," the uttering of this sentence takes time. Not as much time as playing the Goldberg Variations, but still, it unfolds from the first word to the last over the course of 3 seconds or so. My hearing of this utterance, analogously to Soteriou's example, takes more or less the same amount of time as it takes vocalizing it. Crucially, my understanding of this utterance starts when I hear the first word and continues as I hear subsequent words. Just as the critic does not wait with analyzing the Goldberg Variations till the accomplishments of *playing* and *hearing the Goldberg Variations* are finished, our cognitive systems do not "wait" with understanding an utterance till the end of accomplishments of *uttering* and *hearing this utterance*.

Obviously, the accomplishment *understanding an utterance* is intimately related with both state of understanding and an achievement of understanding. The state of understanding an utterance is the *telos* towards which the accomplishment progresses. The achievement, as indicated by Longworth, is the boundary between the accomplishment and the first moment of being in the state of understanding; an instantaneous episode of coming to state-understanding. Stout observes:

> In English the form of words used to refer to an achievement is very often the same as the form of words used to refer to a process. The phrase 'my crossing the road' serves both purposes. This is an instance of the familiar process/product ambiguity. We might say that the writing that was happening in the exam hall resulted in the writing that could be found later on the exam papers. The word 'writing' is functioning differently in its two occurrences here, first to refer to a process and then to refer to the product of that process. (Stout, 2018, 212)

Arguably, in contemporary English, "understanding" is almost never used as progressive. We use expressions like "Anne is crossing the road." or "Anne is analyzing the data." but "Anne is understanding the data." does not sound right. In this context, it is much more natural to say, e.g., "Anne is trying to understand the data." Is it a problem for my project of characterizing linguistic understanding as an accomplishment?

I do not think it is. Even if the progressive use of "understand" is not common, it is not unprecedented.[11] Moreover, the frequency and context in which different words are used in the progressive undergo diachronic changes and are subject to dialectical differences (Smitterberg, 2005; Levin, 2013). We encounter at least a trace of progressiveness in constructions like: "If I'm understanding the results correctly…" This phrase can be interpreted as conveying both that I am in a state of understanding the results in some way (and I am not sure whether this way is the right way) and that I am still in the process of understanding these results (which, hopefully, will

[10] Certainly, analyzing the *recording of Goldberg Variations* may continue long after the *playing* and *hearing* are finished (in fact, it may take place over the course of multiple *playings* and *hearings*), while *understanding an utterance* typically finishes right after the *vocalizing* and *hearing*.

[11] See Jørgensen (1991) for a discussion of progressive uses of predominantly stative verbs including "understand."

culminate with understanding them correctly). To make the process sense even more explicit, imagine that I am reading the results as they are being printed page-by-page. After reading just a few pages, I observe certain tendencies and immediately report: "If I'm understanding the results correctly, the asteroid is going to hit us in 16 hours." I am still in the process of understanding these results, but my current best guess is that the results indicate that the asteroid will hit us in 16 hours. I take the process of linguistic understanding to be analogous to the process of understanding the results as they are being printed page-by-page. Again, following Vendler, my goal is "not to give rules about how to use certain terms but to suggest a way of describing the use of those terms." (Vendler, 1957, 144).

But would it not be better to simply drop the controversial progressive use of "understand" and use *coming to understand an utterance* when speaking about accomplishments? The problem with using *coming to understand* is that it suggests that understanding "is still entirely in the future" (Galton, 2018, 55). However, as my comparison with *analyzing the Goldberg Variations* was intended to demonstrate, and as I will further argue by appeal to empirical research in the next section, when we listen to someone vocalizing an utterance, understanding is not entirely in the future. It is in progress from the get-go and unfolds over time, simultaneously with *vocalizing* and *hearing*. Word after word there is more and more *understanding an utterance* (an accomplishment) leading to *understanding an utterance* (an achievement), which, in turn, demarcates the first moment of *understanding an utterance* (a state). Vendler is on point: "verbs that call for two or more time schemata will provide the most interesting instances of conceptual divergence." (1957, 144).[12]

Finally, can we go one step further and think about the use of *linguistic understanding* as an activity? I think we can. Here is Rob Boffard describing his experience with a speed-reading app in a Guardian article: "I started off at a relatively mild 350wpm, but soon graduated to a cruising altitude of 600wpm. And I was understanding all of it." (Boffard, 2014, Sptember 22). The more *reading* (activity), the more *understanding* (activity). After reading (and understanding what he was reading) for ten minutes, Boffard could have read (and understood what he was reading) for another five minutes, and there would be fifteen minutes of reading with understanding, just like we expect from activities.

In this section I described various uses of *linguistic understanding*. Using *linguistic understanding* as denoting a state or an achievement, corresponding to the ways in which the verb "understand" is typically used in contemporary English, was already common in the debate. However, I argued that we can also use *understanding* as denoting a process (an accomplishment or an activity). From the philosophical point of view, there are good reasons to do so. The temporal structure of the process of understanding overlaps with temporal structures of other processes, both activities (e.g., *reading*) and accomplishments (e.g., *hearing an utterance*).

[12] Notice also, that (in the accomplishment sense) the sentence "Anne was understanding what Tom was saying" is consistent with its never being the case that she understood what Tom was saying. If Tom was saying "The results clearly indicate that the asteroid is going to hit us in 16 hours." but he was wrong, and the asteroid hit right after he vocalized the word "asteroid" annihilating all life on earth, Anne was understanding what Tom was saying but she never understood it. We can say that she understood a part of it, but we can say the same thing about Alice who, when the asteroid hit, was *building a house* (a classical example of an accomplishment). Alice never built the house (achievement), even though she built a part of it, for example, just the walls.

3.2 The process of understanding in psycholinguistics

3.2.1 Incrementality

Above, I have argued that, given its temporal structure, there are good reasons to think about linguistic understanding as a process. In this section, I will try to demonstrate that it perfectly fits the way in which *linguistic understanding* is being characterized by empirical language sciences.

Psycholinguistics commonly characterizes linguistic understanding as unfolding over time. A specific term has even been applied to the type of progressiveness in language comprehension. Psycholinguists say that language comprehension is "incremental," i.e., new unfolding input is processed by the receiver chunk by chunk. This chimes well with the turn-taking data. Given the very short gaps between turns, it is simply impossible for a receiver's cognitive system to "wait" until the speaker's turn has finished and only then to "trigger" the comprehension process and produce a response. However, there are also other phenomena that support the claim regarding the incrementality of language comprehension. One of the most robust and widely studied is "garden pathing", first identified by Thomas Bever (1970). There is a large and fast-growing body of literature on garden-path processing of written text comprehension; however, the phenomenon has also been extensively studied in the context of speech processing (Frazier and Fodor, 1978; Kjelgaard and Speer, 1999; Steinhauer et al., 1999; Schafer et al., 2000; Pauker et al., 2011). Take the sentence "The old man the boat.". When receivers hear or read the full sentence, they realize that the most likely interpretation of the sentence's meaning they had built after hearing or reading only "The old man" (i.e., that "the old man" is a noun phrase) is incorrect. As might be expected, this effect is blocked or reduced by appropriate prosodic markers (in speech) and punctuation (in writing). Nevertheless, the very fact that the phenomenon occurs in some utterances is sufficient to show that language comprehension is incremental. If no interpretation had been formed before the full utterance was heard, receivers would not be surprised by how the utterance actually unfolds.

Another important argument in favor of the incrementality of language understanding comes from cross-modal eye-tracking experiments (Altmann and Kamide, 1999; Kaiser and Trueswell, 2004; Tanenhaus et al., 1995). In their influential study, Altmann and Kamide (1999) recorded the eye movements of participants presented with pictures while simultaneously listening to recorded sentences. The pictures were of objects with clearly differentiated properties, e.g., a boy, a cake, a car, a ball, and a toy train. The participants heard two types of sentences, e.g., "The boy will move the cake." and "The boy will eat the cake." Note that in the second sentence, only one of the objects from the example picture could be referred to as the object the boy was going to eat to form an utterance compatible with common sense (only the cake is eatable). As expected, when there was only one candidate object of a given verb, as in the second sentence above, the probability of the participants looking at that object was significantly increased. Altmann and Kamide conclude that: "information extracted at the verb can be used to guide eye movements to whichever object in the visual context satisfies the selectional requirement of the verb. This guidance is initiated before the linguistic expression corresponding to the verb's direct object is encountered" (p. 258).

One could cite other similar phenomena, but I will follow the psycholinguistic mainstream in taking garden paths, turn-taking, and cross-modal eye tracking to constitute a sufficient argument for the incrementality of language comprehension. Just as suggested above, understanding does not happen after *hearing* or *reading* but happens simultaneously with them.

3.2.2 Prediction

There is another feature of linguistic understanding which speaks to its processual character. It is commonly assumed in psycholinguistics that understanding is not only incremental but also predictive. Understanding an utterance consists of constant formation and updating of predictions about what comes next; how the utterance will unfold and how it will end. Kuperberg and Jaeger (2016) go as far as to state that it is "logically impossible" to reject the claim that language comprehension is predictive in light of some robust linguistic phenomena. Two of those, garden-paths and turn-taking, I have already discussed.[13] Additional evidence is provided by e.g., studies of eye movements during the processing of written texts. It has been repeatedly shown (see references in Kuperberg and Jaeger (2016)) that readers fixate more on unpredictable words and less on words that fit in well with the context (i.e., are easier to predict). For example, in a study conducted by Rayner and Well (1996), participants read sentences containing highly, moderately and weakly constrained target words while their eye movements were recorded. The sentences were taken from an earlier study by Schwanenflugel (1986), who asked participants to list up to three possible ways of completing a given sentence. For instance, for the beginning of a sentence: "The woman took the warm cake out of the ...", the word "oven" was produced more than ninety percent of the time; the word "stove", slightly more than thirty percent; and the word "pantry" less than five percent of the time. Therefore, "oven" is highly constrained, "stove" is moderately constrained, and "pantry" weakly constrained completion of the target sentence. Rayner and Well observed that even though there were no significant differences in gaze-fixation times while reading highly and moderately constrained target words, participants fixated for significantly longer on weakly constrained target words. This indicates that readers constantly predict what word or words will most likely appear next in a given context. Prolonged eye fixation on the unpredictable words reveals the increase in processing difficulty for unpredictable completions.

Let us briefly compare this with the asteroid case discussed in Section 3.1. When I reported that the asteroid is likely to hit us in 16 hours, I was still in the process of understanding the results (subsequent pages containing these results were still being printed), but my best guess about the results was that they indicate that the asteroid is going to hit us in 16 hours. This is similar to what we do when processing linguistic input. Imagine that I answer, "I'd like salad" 0.2 s after hearing "What would you like to have for lunch?" To answer so fast, I must have predicted (given the context, including communicative situation, my knowledge of the interlocutor, etc.), that the question concerns lunch *before* the word "lunch" was vocalized. The time pressure of the efficient linguistic communication (not entirely unlike the time pressure of an upcoming Armageddon)

[13] Notice that the garden-path data is complementary to the turn-taking data. Fast turn-taking alone could result from interlocutors being simply uninterested in what others have to say and merely waiting for their chance to speak. Garden path data, however, proves that hearers build interpretations of what speakers say based on fragments of utterances which they have already perceived. This is why they are surprised when the actual continuations of utterances do not meet their predictions.

forces us to react to an expected result of the process of understanding before the process is completed.

However, the process of *understanding the utterance* does not finish when the process of *uttering the response* begins. This is demonstrated by the fact that when the process of *understanding an utterance* is completed, and the final state of *understanding of this utterance* achieved, we may realize that our prediction has been mistaken. In such a case there are two possible scenarios. Either we have already started to vocalize an inappropriate response and thus we follow it with a repair such as: "I'm sorry, I thought that you were asking about lunch."; or, given enough time and sufficiently keen reflexes, we might block the process of uttering the response and reconsider our own turn. The necessity to repair one's own initial reply, or encountering a highly unexpected utterance, will typically result in a considerably longer gap between subsequent turns (Sacks et al., 1974; Kendrick and Torreira, 2015).

3.2.3 Illocutionary force

There is one more important lesson regarding linguistic understanding that we can learn from the turn-taking studies if we accept LU and extend our characterization of linguistic understanding to account for its processual aspect. Of course, it is a matter of what one's goals are in formulating a theory of linguistic understanding; but it seems that we cannot talk of full understanding of a given utterance unless we mean both understanding of its content and its illocutionary force. Interestingly, this observation finds support in empirical research on turn-taking. According to conversation analysis (Schegloff, 2007; Levinson, 2013), it is impossible to navigate the unfathomable environment of everyday language use and produce subsequent turns with such speed and accuracy, unless we deploy specific strategies and heuristics. One of the most important of these is action ascription.[14] Most discourses are logically organized around so-called "adjacency pairs". For example, if one speaker asks a question, it is highly probable that the turn of the interlocutor will consist of an answer; offers trigger acceptances and rejections; complaints trigger explanations; etc. In general:

> The challenge for participants, then, is to assign at least one major action to a turn they have only heard part of so far. But to do this, they must have parsed what they have heard and understood its grammar well enough to predict both the content and its structure. . . (Levinson, 2013, 103)

[14] Even though the way that conversation analysis characterizes linguistic actions is different from the way philosophy of language typically characterizes speech acts, I take the two enterprises to be complementary. The primary interest of conversation analysis is a statistical classification of countless patterns observable in human linguistic activity (Schegloff, 1992, 2007). Speech act theories, on the other hand, are primarily concerned with constitutive features (either conventional, intentional, functional, expressive or normative) of selected speech acts. Nevertheless, even though the list of speech acts studied in philosophy is shorter than the list of linguistic actions studied in conversation analysis, there can be little doubt that whatever work has already been done in relation to, e.g., asserting, asking, promising, etc., is but the tip of the iceberg of all the speech act types that figure in our everyday linguistic activity. More than anything, it might be a matter of fine-grainedness; maybe speech act theory will come up with a shorter list of more general speech act types that could then be used as scaffolding for the classification of the myriads of subtypes of linguistic actions. As Austin himself already admitted in his foundational work (1962): "I am not suggesting that this is a clearly defined class by any means" (p. 99).

The following picture emerges. In spoken linguistic exchanges understanding begins immediately after the hearer's cognitive system recognizes perceived sounds as linguistic input.[15] The first important task of the hearer's cognitive system is to obtain sufficient information on the grammatical and lexical features of the utterance to classify it as performing a particular speech act. This action ascription, together with background knowledge, guides further processing of the utterance. For example, after hearing "Would you ...?", the comprehension system of an English speaker might initially classify the input as performing the action of asking and use this "assumption" to predict a set of possible grammatical continuations. Obviously, the initial classification is not written in stone. An utterance of "Would you be so kind as to close the window?" has the grammatical structure of a question, but emerging from the mouth of an English speaker would typically perform the action of requesting.

4. Doing away with the process?

In the previous section, I proposed a definition that outlines the relations between different senses of linguistic understanding (LU) and argued that it is appropriate to characterize linguistic understanding as a process. In this section, I will respond to a possible objection according to which we can do away with the process-talk in the context of understanding.

The following worry may arise. Maybe what I am characterizing as a process of understanding is actually just a sequence of states of understanding? For example, when understanding an utterance "What would you like to have for lunch?" a hearer goes through a succession of mental states representing subsequent fractions of the utterance as it unfolds (e.g., "What…", "What would....", "What would you…", etc.).[16] Thus, we can account for the dynamic aspect of linguistic understanding without characterizing understanding as a process.

I think that this worry is ill-grounded. No doubt that the process I characterize can be alternatively modeled as a succession of states. Plausibly, every cognitive process can be conceptualized as a succession of states leading to its final output. Brian O'Shaughnessy (2000) speaks about it when he contrasts non-experiential processes (such as *forgetting*) with experiential processes (such as *hearing*). While experiential processes are "of necessity constituted of nothing but process-parts" because "[e]verything experiential is of necessity in flux." (O'Shaughnessy, 2000, 49), non-experiential processes are constituted of both: state-parts and process-parts. O'Shaughnessy explains:

[15] As I have suggested above, the process of understanding is involuntary. It is widely recognized that linguistic understanding is not an intentional action (Hunter, 1998) and that it is mostly performed by our cognitive system "sub-personally" and independently of our will (Pettit, 2010). We cannot prevent our understanding of a sentence spoken in a language we understand in the dispositional sense, once we hear it. We cannot decide not to trigger the understanding process. Moreover, even though there is an extensive range of actions that we can perform to improve our understanding (in some cases we can facilitate it by, e.g., repeating an exceptionally complicated sentence aloud or writing it down and parsing it), the occasions on which we actually do so constitute but a tiny portion of our omnipresent linguistic activity. Typically, the understanding process initiates automatically, compulsorily and does not require any personal-level action (see, e.g., Marslen-Wilson and Tyler, 1981; Fodor, 1983).

[16] For simplicity I assume here that the utterance is represented in the subsequent states word-by-word. Competing models, e.g., phoneme-by-phoneme, are also possible.

> ….just as being at position p is one thing and being at p and moving is another, so being in a certain memory-state at time t is one thing and being in that state and forgetting some of that content is another. …all that is needed for the occurrence of the requisite process is a continuity across time of states which stand to one another in some requisite relation. With this in mind I say, that when the nonexperiential process of forgetting takes processive form, that process is constituted out of memory-states. (O'Shaughnessy, 2000, 46)

Understanding an utterance is a non-experiential process which overlaps with the experiential process of *hearing an utterance*.[17] Both of them are continuous and "exist by unfolding from one time to another" (Crowther, 2009, 5). The fact that we can point at state-parts constituting the process of *understanding* does not make our characterization of understanding as a process redundant or false.

In the existent debate about linguistic understanding, most authors focused on states of understanding of the whole utterance, i.e. the final outputs of the process of *understanding an utterance*. There is a good reason to pay special attention to such states. As highlighted by Longworth: "What we seek in an account of state-understanding is an account of how such states can play a role in ordinary psychology, how occupying them can impact on the rational development of one's cognitive economy." (Longworth, 2008a, 51-52). States of understanding interact with various other cognitive states of the subject and thus constitute the link between our linguistic activity and other things we do and believe. For example, I go to the kitchen because I believe you said there is carrot cake in the kitchen; I have a desire to have a piece of carrot cake; I know that the kitchen is close enough for me to reach in just couple of seconds; etc. In principle, there is no reason why theories of states of understanding could not be extended to give an account of state-parts of the process of understanding. However, even equipped with a theory of states of partial understanding leading up to full understanding, we should not resign from describing *understanding an utterance* as a process. Like other non-experiential processes, linguistic understanding "is susceptible of two different analyses, according as different analytical agencies are brought to bear upon it" (O'Shaughnessy, 2000, 45).

5. The "process first" approach to linguistic understanding

My main task in this paper was to offer a theory of linguistic understanding that would help us account for the dynamic aspect of everyday linguistic interactions; especially the phenomenon of fast turn-taking. I claimed that to do this we should focus on the nature of the process of understanding. However, since philosophers used to focus, up to this point, mostly on the topic of states of understanding, one might have an impression that my discussion in this essay goes orthogonal to the philosophical debate. There are two things to be said about this worry.

Firstly, even though I appreciate and share the interest of other philosophers in the problem of the nature and function of states of understanding, I suggest expanding our attention to other aspects of linguistic understanding. Human linguistic activity is a very complex and multidimensional phenomenon. One of its most striking features is the speed with which we are able to exchange information during everyday conversations. We cannot explain this feature unless

[17] For more about the nature of experiential processes see Soteriou (2013, 2018).

we take a closer look at the nature of linguistic processing. Even if typically studied using empirical methods, this process may be of interest to anyone investigating the phenomenon of language use.

Secondly, the existent theories are more or less in agreement regarding a few most important criteria a state has to meet to qualify as a state of understanding. Firstly, it has to be a representational state; representing at least the content (and, possibly, also the illocutionary force) of a given utterance. Secondly, it has to be a personal-level state, i.e., a state consciously accessible to the hearer. Finally, it has to be a state that can interact with other personal-level representational states of the subject. These are all necessary conditions for something being a state of understanding. But how can we tell which among the competitive theories is the right one?

LU outlines how the different elements of linguistic understanding: the process, the state, and the disposition to understand a given language, are related to each other. States of understanding are generated by the process of understanding; they are the end products towards which the accomplishments *understanding an utterance* unfold over time. If this is the case, investigating the nature of the process might teach us a lesson about the nature of states.

To illustrate it with an example. It is a central point of Longworth's (2018) theory (in contrast with, e.g., Fricker's (2003) theory) that states of understanding represent the content of the utterance directly (they are attitudes towards the content of the form "p"), and not *as uttered by a given speaker* (i.e., attitudes towards the content of the form "that S said that p" where S represents a speaker). However, since both these alternatives are compatible with the criteria enumerated above, we might be in a better position to say which one is right if we take a closer look at the mechanisms involved in language processing. For example, the data enumerated by Polka and Nazzi (2018) suggest that:

> … although the "who" and "what" sides of the signal are distinct, their processing is functionally integrated, as suggested by research on adults. For example, adults are better at learning to recognize talkers in their native language than in an L2, and dyslexics show impairments in both native language talker recognition and phonological processing… (Polka and Nazzi, 2018, 759)

If the processing of two types of information is in fact integrated, it is not unlikely that it gives rise to a richer type of representation: "that S said that p". Longworth suggests that we should not expect, e.g., small children who seem perfectly able to understand simple utterances, to have conceptual resources necessary to represent content involving concepts such as BEING SAID. But here again the discussion is very much open given that the research on speech processing suggests that:

> Infants successfully extract information about who is talking early in development (the "who" side of the message), well before they begin to sort out the "what" side of the message, that is, the referential content of words. (Polka and Nazzi, 2018, 758)

I am not claiming that this data alone shows that Longworth is mistaken[18] or constitutes conclusive evidence for any particular stance in the debate. It would require at least a separate paper.[19] What I want to illustrate is that if we want to choose between different types of representation performing a given task equally well, it might be beneficial to take a look at the way they are actually produced.

6. Concluding remarks

The goal of this paper was to offer an account of the process of linguistic understanding. Only by focusing on the process sense of understanding, we are able to capture its dynamic and temporal profile. Just like other processes, linguistic understanding unfolds over time.

Crucially, this does not imply that focusing on the state of understanding or the achievement of understanding is inappropriate or misguided. Linguistic understanding is a multidimensional phenomenon and "linguistic understanding" – a capacious notion. As outlined in LU, different aspects of this phenomenon are closely intertwined. My goal was to illuminate the one aspect relatively absent from the philosophical debate to this date. I hope that this will contribute to our gaining an even better grasp on this complex yet fascinating topic in the future.[20]

References

Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition*, 73 (3), 247-264. doi: 0.1016S0010-0277(99)00059 -1

Austin, J. L. (1962). *How to do things with words*. Clarendon Press.

Bach, K. (1994). Conversational impliciture. *Mind and Language*, 9(2), 124–162. doi: 10.1111/j.1468-0017.1994.tb00220.x

Bever, T. (1970). The cognitive basis for linguistic structure. In J. R. Hayes (Ed.), *Cognition and the development of language*. New York: Wiley.

Boeckx, C., & Benítez-Burraco, A. (2014). The shape of the human language-ready brain. *Frontiers in Psychology*, 5, 282. doi: 10.3389/ fpsyg.2014.00282

Boffard, R. (2014, September 22). Can you read a novel in three hours? *The Guardian*. Retrieved from http://www.theguardian.com/books/booksblog/2014/sep/22/speed-reading-app-spritz-read-novel-three-hours

[18] Longworth's account of linguistic understanding is very elaborate (Longworth, 2008a, 2008b, 2010, 2018), and I could not hope to address all its subtleties in the present paper. Here, my only goal is to suggest that by paying special attention to the process by which the states of understanding are generated, we might collect additional data worth taking into consideration while investigating the nature of these states.

[19] I discuss these issues in Author XXX.

[20] I am very grateful to Derek Ball, Liz Camp, Mark Jary, Aleksandra Krogulska, Palle Leth, Manolo Martínez, Michele Palmira, Paul Pietroski, Krzysztof Posłajko, Adam Sennet, and Abel Suñé for helpful comments, questions, and objections. Special thanks to Bartłomiej Czajka, Manuel García- Carpintero, Grzegorz Gaszczyk, and Josep Macià for numerous detailed conversations on the topics of this paper, and an anonymous referee for this journal for excellent suggestions.

Campbell, J. (1982). Knowledge and understanding. *Philosophical Quarterly,* 32(126), 17–34. doi: 10.2307/2218998

Crowther, T. (2009). Watching, Sight, and the Temporal Shape of Perceptual Activity. *Philosophical Review*, 118(1), 1–27. https://doi.org/10.1215/00318108-2008-027

Crowther, T. (2011). The Matter of Events. *Review of Metaphysics*, 65(1), 3–39.

Crowther, T. (2018). Processes as Continuants and Process as Stuff. In R. Stout (Ed.), *Process, Action, and Experience* (pp. 58–81). Oxford University Press.

Davies, M. (1989). Tacit knowledge and subdoxastic states. In A. George (Ed.), *Reflections on Chomsky*. Blackwell.

Dummett, M. A. E. (1993). *The seas of language*. Oxford University Press.

Evans, G. (1982). *The varieties of reference*. Oxford University Press.

Fodor, J. A. (1983). *The modularity of mind*. MIT press.

Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, 6(4), 291–325. doi: 10.1016/0010-0277(78)90002 -1

Fricker, E. (2003). Understanding and knowledge of what is said. In A. Barber (Ed.), *Epistemology of language* (pp. 325–66). Oxford University Press.

Friederici, A. D. (2012). The cortical language circuit: from auditory perception to sentence comprehension. *Trends in Cognitive Sciences*, 16 (5), 262–268. doi: 10.1016/j.tics.2012.04.001

Galton, A. (2018). Processes as Patterns of Occurance. In R. Stout (Ed.), *Process, Action, and Experience* (pp. 41–57). Oxford University Press.

Garrod, S., & Pickering, M. J. (2015). The use of content and timing to predict turn transitions. *Frontiers in Psychology*, 6, 751. doi: 10 .3389/fpsyg.2015.00751

Hauser, M., Chomsky, N., & Fitch, T. W. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, 298(22), 1569–1579. doi: 10.1126/science.298.5598.1569

Heck, R. (1995). The sense of communication. *Mind*, 104(413), 79–106. doi: 10.1093/mind/104.413.79

Higginbotham, J. (1992). Truth and understanding. *Philosophical Studies*, 65(1-2), 3–16. doi: 10.1007/BF00571313

Hunter, D. (1998). Understanding and belief. *Philosophy and Phenomenological Research*, 58(3), 559–580. doi: 10.2307/2653756

Jackendoff, R. (2002). What's in the lexicon? In S. Nooteboom, F. Weerman, & F. Wijnen (Eds.), *Storage and computation in the language faculty* (pp. 23– 58). Dordrecht: Springer Netherlands. doi: 10.1007/978-94-010-0355-1_2

Jackendoff, R., & Pinker, S. (2005). The faculty of language: What's special about it? *Cognition*, 95(2), 201–236. doi: 10.1016/j.cognition.2004.08.004

Jørgensen, E. (1991). The progressive tenses and the so-called 'non-conclusive' verbs. *English Studies*, 72(2), 173–182. https://doi.org/10.1080/00138389108598743

Kaiser, E., & Trueswell, J. C. (2004). The role of discourse context in the processing of a flexible word-order language. *Cognition*, 94(2), 113 - 147. doi: https://doi.org/10.1016/j.cognition.2004.01.002

Kendrick, K. H. (2015). The intersection of turn-taking and repair: the timing of other-initiations of repair in conversation. *Frontiers in Psychology*, 6, 250. doi: 10.3389/fpsyg.2015.00250

Kendrick, K. H., & Torreira, F. (2015). The timing and construction of preference: A quantitative study. *Discourse Processes*, 52 (4), 255-289. doi: 10.1080/0163853X.2014.955997

Kjelgaard, M. M., & Speer, S. R. (1999). Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. *Journal of Memory and Language*, 40(2), 153–194. doi: 10.1006/jmla.1998.2620

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, 31(1), 32-59. doi: 10.1080/23273798.2015.1102299

Levin, M. (2013). The progressive verb in modern American English. In B. Aarts, J. Close, G. Leech, & S. Wallis (Eds.), *The Verb Phrase in English* (pp. 187–216). Cambridge University Press.

Levinson, S. C. (2013). Action formation and ascription. In Sidnell, Jack and Stivers, Tanya (Ed.), *The handbook of conversation analysis* (pp. 103–130). Wiley-Blackwell.

Levinson, S. C. (2016). Turn-taking in human communication – origins and implications for language processing. *Trends in Cognitive Sciences*, 20(1), 6– 14. doi: 10.1016/j.tics.2015.10.010

Levinson, S. C., & Torreira, F. (2015). Timing in turn-taking and its implications for processing models of language. *Frontiers in Psychology*, 6, 731. doi: 10.3389/fpsyg.2015.00731

Longworth, G. (2008a). Linguistic Understanding and Knowledge. *Nous*, *42*(1), 50–79. https://doi.org/10.1111/j.1468-0068.2007.00673.x

Longworth, G. (2008b). Comprehending speech. *Philosophical Perspectives*, 22(1), 339–373. https://doi.org/10.1111/j.1520-8583.2008.00151.x

Longworth, G. (2010). Some models of linguistic understanding. *The Baltic International Yearbook*, 5(1), 7. doi: 10.4148/biyclc.v5i0.284

Longworth, G. (2018). Understanding what was said. *Synthese*, 195(2), 815–834. doi: 10.1007/s11229-016-1243-5

Marslen-Wilson, W., & Tyler, L. K. (1981). Central processes in speech understanding. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 295(1077), 317–332. https://doi.org/10.1098/rstb.1981.0143

Millikan, R. G. (1984). *Language, thought and other biological categories*. MIT Press.

Millikan, R. G. (2004). *Varieties of meaning*. MIT Press.

Mourelatos, A. P. D. (1978). Events, processes, and states. *Linguistics and Philosophy*, 2(3), 415–434. doi: 10.1007/BF00149015

O'Shaughnessy, B. (2000). *Consciousness and the world*. Oxford University Press.

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007, 12). Where do you know what you know? the representation of semantic knowledge in the human brain. *Nat Rev Neurosci*, 8 (12), 976–987. doi: 10.1038/nrn2277

Pauker, E., Itzhak, I., Baum, S. R., & Steinhauer, K. (2011). Effects of cooperating and conflicting prosody in spoken english garden path sentences: ERP evidence for the boundary deletion hypothesis. *Journal of Cognitive Neuroscience*, 23(10), 2731-2751. doi: 10.1162/jocn.2011.21610

Pettit, D. (2010). On the epistemology and psychology of speech comprehension. *The Baltic International Yearbook of Cognition, Logic and Communication*, 5(1), 9. doi: 10.4148/biyclc.v5i0.286

Pinker, S., & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences*, 13 (4), 707-727. doi: 10.1017/S0140525X00081061

Polka, L., & Nazzi, T. (2018). Interacting processes and developmental biases allow learners to crack the "what" code and the "who" code in spoken language. *Applied Psycholinguistics*, *39*(4), 757–761. https://doi.org/10.1017/S0142716418000218

Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review*, 3 (4), 504–509. doi: 10.3758/BF03214555

Rothstein, S. (2004). *Structuring events: A study in the semantics of aspect*. Wiley.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50(4), 696-735. doi: 10.1016/B978-0-12-623550-0.50008-2

Schafer, A. J., Speer, S. R., Warren, P., & White, S. D. (2000). Intonational disambiguation in sentence production and comprehension. *Journal of psycholinguistic research*, 29(2), 169–182. doi: 10.1023/A:1005192911512

Schegloff, E. A. (1992). To Searle on conversation: A note in return. In J. R. Searle, H. Parret, & J. Verschueren (Eds.), *(On) Searle on conversation* (pp. 113–128).

Schegloff, E. A. (2007). *Sequence organization in interaction: Volume 1: A primer in conversation analysis*. Cambridge University Press.

Schwanenflugel, P. J. (1986). Completion norms for final words of sentences using a multiple production measure. *Behavior Research Methods*, 18(4), 363–371. doi: 10.3758/BF03204419

Smitterberg, E. (2005). *The progressive in 19th-century English: A process of integration*. Rodopi.

Soteriou, M. (2013). *The mind's construction: The ontology of mind and mental action*. Oxford University Press.

Soteriou, M. (2018). Experience, Process, Continuity, and Boundary. In R. Stout (Ed.), *Process, Action, and Experience* (pp. 82–101). Oxford University Press

Steinhauer, K., Alter, K., & Friederici, A. D. (1999). Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature neuroscience*, 2(2), 191–196. doi: 10.1038/5757

Steward, H. (1997). *The ontology of mind: Events, processes, and states*. Oxford University Press.

Steward, H. (2012). Actions as processes. *Philosophical Perspectives*, 26(1), 373–388. doi: 10.1111/phpe.12008

Steward, H. (2018). Ocurrent States. In R. Stout (Ed.), *Process, Action, and Experience* (pp. 102–119). Oxford University Press.

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., . . . others (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences*, 106 (26), 10587–10592. doi: 10.1073/pnas.0903616106

Stout, R. (2018). Ballistic Action. In R. Stout (Ed.), *Process, Action, and Experience* (pp. 210–228). Oxford University Press.

Tanenhaus, M., Spivey-Knowlton, M., Eberhard, K., & Sedivy, J. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science* , 268 (5217), 1632–1634. doi: 10.1126/science.7777863

Vendler, Z. (1957). Verbs and times. *Philosophical Review*, 66 (2), 143–160. doi: 10.2307/2182371

Verkuyl, H. J. (1972). *On the compositional nature of the aspects*. D. Reidel Publishing Company.