

## Grounds, Convention, and the Metaphysics of Linguistic Tokens

BRIAN EPSTEIN

*Croatian Journal of Philosophy* 9,  
No. 25 (2009).

---

*My aim in this paper is to discuss a metaphysical framework within which to understand “standard linguistic entities” (SLEs), such as words, sentences, phonemes, and other entities routinely employed in linguistic theory. In doing so, I aim to defuse certain kinds of skepticism, challenge convention-based accounts of SLEs, and present a series of distinctions for better understanding what the various accounts of SLEs do and do not accomplish.*

In the last few years, a number of philosophers have debated the nature of what Georges Rey has called “Standard Linguistic Entities” (SLEs), i.e., words, sentences, phonemes, noun phrases, and other entities routinely employed in linguistic theory.<sup>1</sup> Positive theorists of SLEs commonly take themselves to be in the business of giving the conditions for a physical token, such as an acoustic blast or stream, to be an instance of a particular standard linguistic entity. Skeptics deny that sets of tokens are interestingly or relevantly unified into types or kinds.

My aim in this paper is to clarify the metaphysical landscape within which to understand SLEs. In doing so, I aim to defuse at least certain kinds

---

\* I am grateful to the participants in the 2008 Dubrovnik conference in philosophy of linguistics, and to Guy Longworth and Georges Rey, for their comments.

<sup>1</sup> Barber 2006, Devitt 2006b, Devitt 2006a, Rey 2006c, Rey 2006a, Rey 2006b, Barber 2008, Collins 2008, Rey 2008. Earlier discussions related to this topic include Millikan 1984, Bromberger 1989, Kaplan 1990, Richard 1990, Cappelen 1999, Szabo 1999, Bromberger and Halle 2000.

of skepticism, present an explicit challenge to one positive account of SLEs, and to set up a framework for better understanding what the various accounts of SLEs do and do not accomplish. I begin by applying some existing distinctions in metaphysics to SLEs, and then introduce some new distinctions, in particular between *tokening conditions* and the *grounds* for those conditions. In connection with these distinctions, I consider Herman Cappelen's and Michael Devitt's attempts to ground the individuation of SLEs in convention. Although convention is useful for helping illuminate the distinctions I introduce, I argue that once the metaphysical landscape and the questions being addressed are clarified, it becomes evident that convention-based accounts are not correct.

Understanding the nature of SLEs is interesting in its own right, and is likely to be relevant to linguistic methodology. Another reason the metaphysics of SLEs is important, however, is that it promises to be connected to the metaphysics of other special sciences as well. For working purposes, I will draw on the metaphysics of other special sciences and the metaphysics of ordinary objects to clarify aspects of SLEs. Conversely, getting clear on the metaphysics of SLEs will, I hope, advance our understanding of those other fields.

This paper consists of five parts: (i) briefly review some phenomena pertaining to acoustic streams discussed by Rey; (ii) apply some aspects of the metaphysics of ordinary objects to clarify intrinsic and extrinsic essential properties of objects, and to explain why we might expect sets of acoustic blasts of the same linguistic type to be highly irregular or "gerrymandered"; (iii) reconsider exactly what the question is that accounts of SLEs mean to address, in light of the fact that SLEs appear at different levels in a kind of ladder of abstraction; (iv) introduce the important distinction between tokening conditions and grounds, using the example of convention; and (v) criticize the convention-based answer to the grounding of SLEs.

## 1. Acoustic unruliness and variability

Georges Rey describes a number of different phenomena that show why the acoustic properties of linguistic utterances are not well behaved.<sup>2</sup> First, it is common for there to be no boundaries between linguistic units, such as the break between words in a normal utterance. For instance, the pronunciation of “a tack” is typically indistinguishable from “attack.” Second, there is not a clear correspondence between psychological and acoustic segments. For instance, we do not produce the monosyllabic word ‘bag’ by producing the sequence of phones out of which it is composed. In general, there is no good way to slice up an acoustic stream into tokens of individual phones. Third, there is complexity in establishing physical acoustic types corresponding to phonemes. The phoneme /d/, for instance, is tokened by a different acoustic event depending on the vowel which it precedes. Fourth, there is the interesting phenomenon in which a single sound can be heard as different phonemes, depending on what precedes or follows it in the acoustic stream:

*In normal speech, the difference between “rider” and “writer” is heard as a difference between /d/ and /t/. But in fact the /d/ and /t/ are pronounced identically (as flaps), the difference in how they’re heard being due to the longer sound of the preceding vowel.”<sup>3</sup>*

These are good pieces of evidence for what we might call “acoustic unruliness.” What Rey calls “variability” is slightly different, and is also relevant to the failure of acoustic streams to be well behaved. In Rey’s terminology, variability is difference in pronunciation and articulation by different individuals or by a single individual across contexts. There are not only dialectal sources of variation in how words are pronounced across individuals, but also differences according to age, circumstances, level of inebriation, and so on. This variability makes it unlikely that phonetic features can be typed based on acoustic characteristics at all. In any account

---

<sup>2</sup> Rey 2006b

<sup>3</sup> Rey 2006b, p. 7.

of the typing of linguistic tokens, all of these phenomena have to be dealt with or explained.

## 2. Modeling gerrymandered sets of acoustic blasts

A matter that has been poorly understood in the SLE literature is whether standard linguistic entities should be identified with – i.e., identical to – acoustic phenomena. Alex Barber has recently called this view “acoustic reductionism,”<sup>4</sup> pointing out that Rey endorses it explicitly:

*Part and parcel of this commonplace [i.e., that people produce tokens of words etc.] would be the presumption that these entities can be identified with some sorts of acoustic phenomena, e.g., wave patterns in space and time.*<sup>5</sup>

Barber rejects acoustic reductionism by proposing a non-acoustic set of identity conditions for SLE tokens. I want to begin a step earlier, and at the outset present a picture of SLE-tokens in which they are understood to be constituted by but not identical to acoustic patterns. Below, I will discuss a new and equally important distinction, between the tokening conditions of a type and the grounds for those conditions.

---

<sup>4</sup> Barber 2006

<sup>5</sup> Rey 2006c. Barber also labels Devitt and Sterelny as acoustic reductionists, citing their claim that: “Tokens are datable, placeable parts of the physical world. Thus Nana and her successor, Lulu, are cat tokens... Types, on the other hand, are kind of tokens. Any token can be grouped into many different types... Inscription types and sound types are identifiable by their overt physical characteristics and so we might call them “physical types.”” (Devitt and Sterelny 1999) But it is not clear that Barber’s interpretation of Devitt and Sterelny correctly applies to their view of SLEs as opposed to just inscription and sound types, especially since they say in the following sentence, “Word tokens are also grouped semantically.” Devitt in more recent work has stressed that his picture of linguistic entities is that they are relational, like Australians.

## *2.1 Using a liberal ontology as a working approach*

The nature of SLEs is a question in linguistic metaphysics, and it would be helpful if we could draw on advances in metaphysics to frame the questions pertaining to SLEs. Bringing that work in, however, is hindered by two considerations.

First, many pertinent issues in the metaphysics of ordinary objects remain unresolved. It continues to be debated, for instance, whether a distinction ought to be made between constitution and identity. And supposing the distinction is made, the nature of the constitution relation is debated, as well as the question whether there is a single constitution relation. Further, there is a spectrum of views about what is thereby implied about the individuation of ordinary objects. While some metaphysicians take it to sanction coincident entities, others take a distinction between constitution and identity to be compatible with the failure of coincidence.<sup>6</sup>

Second, it is unclear to what extent the metaphysics of ordinary objects should apply to linguistic entities at all. Some philosophers take these issues to pertain only to “manifest” rather than “scientific” entities. Even if we were to resolve the problems of “manifest linguistic objects,” this may be irrelevant to the determination of entities that are relevant to linguistic theory.

It is not my aim to resolve either of these problems. I will, however, take a position on them for our working purposes. In particular, I adopt a liberal ontology, and try to make sense of SLEs as ordinary objects, without imposing harsh conditions in advance about what kind of object might or might not count as scientific. Still, I only intend to use it as a model, mainly to disambiguate terminology and positions.

## *2.2 A liberal model of objects and its application to SLEs*

In replying to Rey’s skepticism about SLEs, Devitt speaks of them as “essentially relational entities”:

---

<sup>6</sup> I am grateful to Guy Longworth for emphasizing these points. These issues among others are explored in the papers in Rea 1997.

*In general, those who present considerations against the existence of SLEs need to show how those considerations, if valid, would not count equally against the existence of many other essentially relational entities, particularly social entities, that clearly do exist. Rey's considerations do not pass this test.*<sup>7</sup>

An important point of difficulty in these discussions is that it is often unclear whether the entities being referred to are SLE-types or SLE-tokens. When speaking of a SLE as a potentially relational entity, can it make sense to speak of SLE-tokens as relational? Tokens are individual objects, but in what sense can a relational entity count as an individual?

One straightforward interpretation is that the individual objects in question, which may be tokens of SLEs, are acoustic blasts. A particular acoustic blast is a spatiotemporal particular, which may or may not have the property *being a word* or *being a phone*. Some people insist that the possession of such a property must depend only on the intrinsic properties of an acoustic blast, while others propose that the property may be extrinsic or relational.

A different but perfectly consistent way of speaking is to be more generous about what counts as an individual object. It is familiar from the abundant literature on Goliath and Lump1 (Alan Gibbard's names respectively for a particular statue and for the clay out of which it is constituted)<sup>8</sup> that we often speak of objects whose essential properties differ from those of the material of which they are constituted. To set out a model for working purposes, I will make use of a small fragment of Kit Fine's theory of objects in "Things and their Parts." Fine understands an object to be a special sort of composite: given a set of objects  $a_1, a_2, \dots$ , and a relation  $R$  in which the objects stand, there is an object (notated  $a_1, a_2, \dots /R$ ) that is the whole consisting of those objects standing in that relation. A ham sandwich, for instance, consists of two slices of bread and a slice of ham (as  $a_1, a_2$ , and  $a_3$ ),

---

<sup>7</sup> Devitt 2006a

<sup>8</sup> Gibbard 1975

where R is the arrangement  $a_3$  is between  $a_1$  and  $a_2$ . He calls  $a_1, a_2, \dots$  the “matter” of the object, and R the “principle of rigid embodiment” of the object.<sup>9</sup>

Or consider a particular wedding ring. It is a piece of metal; let us suppose that it is made of a large number of gold atoms  $a_1, a_2, \dots, a_n$ . The piece of metal itself has the property *being a wedding ring* (call it property W), yet of course the property is not intrinsic. For a piece of metal to have W depends on various social and historical factors being in place. In contrast, the property *being a ring* (call it R) is intrinsic.<sup>10</sup>

On this approach, there are several objects, all having the same matter: the lump of gold, the ring, the wedding ring, and so on. Using Fine’s notation, let us denote these different objects as the matter paired with the appropriate relation. Let A abbreviate the list of all the gold atoms  $a_1, \dots, a_n$  that constitute each of these coinciding objects. The ring and the wedding ring are, respectively, A/R and A/W. And to denote the lump of gold alone as a fusion of atoms, we can pair the matter with a relation that applies to all objects, such as  $x=x$ . So the lump of gold is A/ $x=x$ .<sup>11</sup> All these objects have the intrinsic property *being a ring* and the extrinsic property *being a wedding*

---

<sup>9</sup> I will only consider the “rigid embodiment,” rather than his theory of “variable embodiments,” for our purposes.

<sup>10</sup> It may be worth pointing out a simple distinction that is not too relevant here, but that may potentially be a source of confusion. There is an important difference between a property being socially generated and its being socially individuated. Even the property R (i.e., *being a ring*) in a sense is socially generated, i.e., there is some historical-social account of why we have the concept or employ the property. R is not necessarily a property of basic physics. We might say that it salient to us because of social considerations. But importantly, R is intrinsic, holding of an object just in case the object satisfies certain structural conditions. Despite its being socially generated, its holding does not depend ontologically on social factors.

<sup>11</sup> Or, to be a lump may instead require spatial contiguity, in which case  $x=x$  should be replaced with a stronger relation.

*ring*. Despite their having the same intrinsic and extrinsic properties,<sup>12</sup> however, the different coinciding objects have different *essential* properties. The property *being a wedding ring*, for instance, is essential to A/W but not to A/R. Thus A/W has an extrinsic essential property. The property *being a wedding ring* requires that certain social and historical relations are in place, and standing in those social and historical relations is essential to that object being what it is.

This model thus gives us a way of talking about individual objects that are extrinsically individuated. (I think this will be a somewhat clearer terminology than to call them “relational.”) Although having such an abundance of objects is ontologically suspect, it helps us to lay out terminology to untangle some of the SLE debates.

So let me now refine some terminology. Standard linguistic entities like *word* and *phoneme*, as well as a particular word like ‘Aristotle’ or a phoneme like /d/ are entities that individual objects are tokens of. I will call these entities *types*,<sup>13</sup> and I will work with a liberal conception of types as well, where a given token may fall under any number of types.<sup>14</sup> Sometimes it will

---

<sup>12</sup> They may not have all the same properties; see Fine 2003.

<sup>13</sup> It might be preferable to call them “kinds,” though that does a certain amount of violence to the notion of a kind as a natural class of entities, and admits any class of entities having a property, however artificial, to count as a kind. Still, calling them kinds would be in line with recent terminology in social ontology, in which such categories as race and gender are referred to as “social kinds” (e.g., Haslanger 1995, and others). In using either term, I do not mean to rule out an account of individual words like David Kaplan’s, in which a word is understood to be an individual, and the particular uses of the word are stages or parts in the continuant. Taking SLEs to be continuants, the same distinction holds, where a temporal part or stage will have an acoustic incarnation but requires extrinsic factors (like certain intentions, in Kaplan’s account), in order to count as a stage.

<sup>14</sup> Since both types and objects are understood liberally, I do not make any strong presumptions about the relation between types and their tokens. I do not assume that a token of a type is individuated by that type, or that being a member of the type is essential to the token.

be preferable to replace talk of SLE-types with talk of properties such as *being a word* or *being the word 'Aristotle'*.<sup>15</sup>

From the above discussion, it is clear why speaking of “tokens” of SLEs is problematic. It is often assumed that uttered tokens of SLEs are just acoustic blasts. But when we speak of a particular token of *statue*, are we referring to Goliath or just to Lump1? Similarly, when we speak of a token of a word, do we refer an extrinsically individuated entity analogous to A/W, or to an entity analogous to A/x=x? If it seems obvious that it is the latter, it might be noticed that individuating acoustic blasts may itself be problematic. Whether treated as objects or events, it is not apparent what their identity conditions or modal properties are, nor how the “matter” is to be filled in, such that the “principle of rigid embodiment” is empty.<sup>16</sup>

To cut through this, let us fall back on reasonable assumptions. Putting aside any difficulty in individuating acoustic blasts, let us take them to be the “matter” out of which SLE-tokens are embodied together with a “principle of rigid embodiment.” Let us also admit extrinsically individuated SLE-tokens, while being sure that we introduce terminology to distinguish such tokens from their matter.

I will use the term ‘token’ generally, so that objects liberally understood can be tokens of a type. Suppose, for instance, that B is an acoustic blast that sounds like a well-formed utterance of ‘Aristotle’, and that R is a potentially extrinsic property that applies to acoustic blasts, such as the property *having been uttered with the intention of producing a token of 'Aristotle'*. Also suppose that B actually has property R. Then there is an object B/R, an object that may be both a token of ‘Aristotle’ and a token of *word*. Following Fine, I will refer to the acoustic blast B as the *matter* of the token.

---

<sup>15</sup> One of the reasons for talking about properties rather than objects is that supervenience and extrinsicness are much clearer in speaking about properties. Another reason is that speaking of properties can make type/token ambiguities more obvious.

<sup>16</sup> Treating acoustic blasts as objects in classical mereology is clearly inadequate, since they involve structural properties as well as material ones. So if tokens must be objects of classical mereology, then acoustic blasts cannot be tokens.

It is important to note the following. Although a broad array of objects may thus be SLE-tokens, we need to be careful in speaking of acoustic blasts themselves as tokens of SLE-types. Let us suppose B/R is indeed a token of the SLE-type *word*. Often it is useful to speak of B itself as standing in some token-like relation to *word*.<sup>17</sup> However, it is important to avoid suggesting that it is in virtue of the intrinsic properties of B alone that *word* is tokened. Therefore, rather than speaking of B as an “acoustic token” of a SLE-type, I will speak of it as an “incarnation” of the SLE-type.<sup>18</sup> “An incarnation” of a type, strictly speaking, is shorthand for “the matter of a token” of the type.

To anyone allergic to this liberal ontology, it may be possible to recast the issue back again to the possession of extrinsic properties by acoustic blasts. But considering the tokens of SLEs to be acoustic blasts so often misleads people into insisting that the typing of those tokens can only depend on acoustic features, that it may be worth it to employ this ontology at least as a way of speaking.

### *2.3 Is gerrymandering reason for skepticism about SLEs?*

Rey’s evidence suggests that if we consider a set of acoustic blasts that are all incarnations of some given word or phone, that set will look gerrymandered<sup>19</sup> rather than being a natural class.

When we try to work out the basis for a set of objects to be grouped into a type, a first step might be to look for commonalities among their intrinsic

---

<sup>17</sup> It may be that if B/R is a token of some type K, then B is also a token of K. Or if Fine 2003 is correct, then it need not be the case. Either way, it can be useful to note that it is the matter of a token of K.

<sup>18</sup> Presenting B as an “acoustic token” of K risks suggesting that K is an acoustic type, when it may only be a type whose tokens all have acoustic incarnations, but whose tokening conditions involve extrinsic as well as acoustic factors.

<sup>19</sup> For non-American readers: the term ‘gerrymandered’ was originally applied to congressional districts that were carved up in baroque and unnatural ways (to some people they looked like salamanders), usually done in the interest of biasing future elections in favor of the party in power.

properties. But this, of course, is not the only way of grouping objects into types. It can in fact be trivial to produce gerrymandered sets of incarnations of some type. Gerrymandered incarnations can arise from a liberal conception of types even with a strict conception of objects. Suppose we have a domain of objects strictly conceived, e.g., individuated as in classical mereology. But suppose there is no restriction on the properties that can be used for sorting them into types. They can, for instance, be classed in types according to their functional role, such as whether they function to turn screws or to transport people around the city. Considering the objects within such a type, we expect the incarnations to be gerrymandered.

Gerrymandered sets of incarnations can equally trivially arise from a liberal conception of objects even with a strict conception of types. We might limit typing very strictly, such as taking objects to belong to the same type only if they have the same essential properties. But taking a domain of objects broadly conceived, such as wedding rings (e.g.,  $A_1/W$ ,  $A_2/W$ ) and engagement rings (e.g.,  $A_3/E$ ,  $A_4/E$ ) and friendship rings (e.g.,  $A_5/F$ ,  $A_6/F$ ), we nonetheless generate gerrymandered sets of incarnations. All the various wedding rings would be in one type, while an engagement ring that was physically indistinguishable from a wedding ring would be in a different type.

Gerrymandering is a reason for skepticism about the existence of types or about the existence of a certain sort of objects only if it is a small part of a much more ambitious argument. As Rey notes, the crucial part of his argument against SLE-types is that no such types play a role in linguistic explanation.<sup>20</sup> Part of his argument seems to be one about causation, insisting that only acoustic features are relevant for the causal effects SLE-tokens have. If correct, which is debatable, this still leaves it open that SLE-types are extrinsic, like the type *screwdriver*. So a different part of the argument seems to be that SLE-types happen as a contingent matter not to be useful in actual linguistic explanation. Arguing these, it seems to me, is an uphill battle. But they do get the place of the gerrymandering evidence correct: gerrymandering

---

<sup>20</sup> Rey 2008

on its own is little reason for skepticism without these additional, far more important, pieces.<sup>21</sup>

### 3. What question are these accounts trying to answer?

There are two ambiguities to resolve in clarifying the intuitive question that accounts of SLEs aim to answer.

First, the question *what are the tokening conditions for a SLE?* can be cashed out either as a question about tokens or about incarnations. An example of the first is: *what are the conditions a particular entity needs to satisfy in order to be a token of the word ‘cat’, or the phone [d]?* And an example of the second is: *what are the conditions for a particular acoustic blast **cat**<sub>1</sub> to be an incarnation of the word ‘cat’, or for a particular acoustic blast **d**<sub>1</sub> to be an incarnation of the phone [d]?* (Expressions in bold and with subscripts are names for spatiotemporal particulars, i.e., names for acoustic blasts.<sup>22</sup>) The good news is that both questions amount to more or less the same thing. A particular acoustic blast **cat**<sub>1</sub> is an incarnation of ‘cat’ just in case there is an entity E such that (i) E is a token of ‘cat’, and (ii) **cat**<sub>1</sub> is the matter of E. So an answer to the incarnation question effectively needs the

---

<sup>21</sup> Where Rey does seem to go astray is in insisting that for SLEs to exist, their linguistically relevant features need to be “entokened in the acoustics.” In Rey 2007 he cites Liberman 1996 for instance, in arguing that if all the relevant phonological information were entokened in the acoustics, then it would be too complex for us to decode it. Yet any account of SLE-types or tokens as extrinsic takes them to have linguistically relevant features that do not show up in the acoustics, just as a functional account of the type *screwdriver* takes there to be screwdrivingly-relevant features that do not show up in the screwdriver-incarnation, such as the properties of the screws which they are designed to turn and the hands that are designed to hold them. As with the gerrymandering data, the Liberman data can be taken as evidence that SLE-type-properties are extrinsic, or that SLE-tokens are extrinsically individuated, or both.

<sup>22</sup> To be more specific, suppose that in presenting this paper on Sept. 10, 2008, I uttered “cat cat dog dog d” from 11:00:00am to 11:00:05am. Then (presuming I timed it correctly), let **cat**<sub>1</sub> refer to the acoustic blast uttered from 11:00:00 to 11:00:01, **cat**<sub>2</sub> to the acoustic blast uttered from 11:00:01 to 11:00:02, and so on.

same response as an answer to the tokening question. Often it will be preferable to state and consider the incarnation question, since is the one that is usually explicitly asked, and it is easy to refer to particular incarnations (such as some acoustic particular **cat<sub>1</sub>** or **d<sub>1</sub>**).

However, there is a second problem that has been largely overlooked. Namely, exactly what types are the ones whose identity conditions we aim to investigate?

Fig. 1

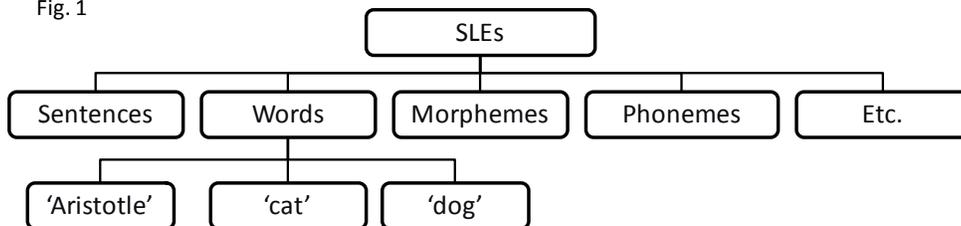


Figure 1 puts a few standard linguistic entities into a hierarchy. Let us focus on the question as to what the incarnation conditions are for a *word* in particular.<sup>23</sup> Considering a particular acoustic blast, **cat<sub>1</sub>**, that blast is an incarnation of the word ‘cat’, but it is also an incarnation of the type *word*, and it is also an incarnation of the type *standard linguistic entity*.<sup>24</sup> Which of

<sup>23</sup> For our purposes, phonemes would actually be a somewhat more complicated case, because of the distinction between phonemes and phones, where *phoneme* is generally taken to denote a type with strictly neural incarnations while *phone* is a type with strictly acoustic incarnations. We could consider the identity conditions of phones, but that seems a narrower inquiry than the question of the identity conditions for SLEs in general. With words we do not have that distinction, so they are a better example of entities that have many kinds of incarnations and represent typical entities employed in linguistic theory.

<sup>24</sup> Again, I do not mean with this terminology to rule out an account of individual words like David Kaplan’s, in which a word is understood to be an individual, and the particular uses of the word are stages or parts in the continuant.

Along with a liberal ontology of tokens, there is a different sense in which we may want to say that the subtypes are members of the higher levels of the hierarchy, e.g., the word ‘Aristotle’ is an instance of *word*. The issue I want to stress is that when we are asking what the conditions are for being an incarnation of a SLE, we need to decide which level

these are we interested in, when we want to understand why gerrymandered sets of acoustic blasts are all incarnations of the same standard linguistic entity? I suggest that the answer is not the one we might first guess.

What many philosophers concerned with the nature of SLEs are attempting to answer, I propose, is the question of the *generic* tokening conditions of *particular* words. To put it in terms of incarnations: we are interested in the conditions for an acoustic blast to be an incarnation of a particular word, rather than to be an incarnation of the type *word*; but the key question is the generic incarnation conditions that hold for all particular words. Let me explain.

First, that we are typically concerned with conditions for particular words. Consider some pairs of acoustic blasts: **cat<sub>1</sub>** and **dog<sub>1</sub>**, versus **cat<sub>1</sub>** and **cat<sub>2</sub>**. The acoustic blasts **cat<sub>1</sub>** and **cat<sub>2</sub>** are incarnations of ‘cat’, while all three of these spatiotemporal particulars are incarnations of words. Compare two different questions: (1) What makes **cat<sub>1</sub>** and **cat<sub>2</sub>** count as incarnations of ‘cat’? And (2) What makes **cat<sub>1</sub>** and **dog<sub>1</sub>** count as incarnations of the type *word*? Even the hopeful physical reductionists would not look for the acoustic commonalities between **cat<sub>1</sub>** and **dog<sub>1</sub>**. Rather, the typing is meant to answer question (1), i.e., to find commonalities between incarnations of ‘cat’. Equally, those who put forward alternatives to physical typing typically seek to provide an answer to (1).

But second, that we are typically concerned with *generic* conditions. It is not exactly right to see the prevailing goal as looking for the conditions for an acoustic blast to count as an incarnation of ‘cat’. The reason is that we expect that there will be the same principle for unifying the incarnations **cat<sub>1</sub>** and **cat<sub>2</sub>** and for unifying two incarnations **dog<sub>1</sub>** and **dog<sub>2</sub>**. We are not looking for the identity conditions for ‘cat’, but for the generic identity conditions that apply across words, for explaining why **cat<sub>1</sub>** and **cat<sub>2</sub>** are tokens of ‘cat’ and why **dog<sub>1</sub>** and **dog<sub>2</sub>** are tokens of ‘dog’ and why **Aristotle<sub>1</sub>** and **Aristotle<sub>2</sub>** are

---

of the hierarchy we are focusing on. For this reason, I will just treat these types as though they can be tokened by the same entities that their subtypes can.

tokens of ‘Aristotle’, and so on. The aim of an account of the incarnation conditions for words, that is, is to answer the question:

(GIC) Given any word W, what are the conditions for an acoustic blast to be an incarnation of W?

I will call this the question of generic incarnation conditions for words. This involves an assumption – I think a correct assumption – that there *is* one generic set of conditions across different words. We can imagine that that is not the case (e.g., that the answer to GIC will be a disjunctive set of conditions, with one set for words starting with the letter ‘a’, a different for words starting with ‘b’, and so on), but it is reasonable to take there to be one answer to this question for all words, and all theorists seem to assume this.

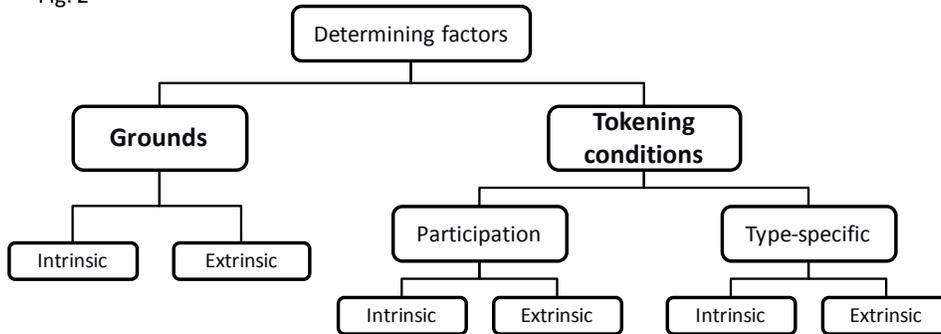
To summarize the key points so far: (1) on our liberal working model, acoustic blasts are the incarnations of tokens of SLEs; (2) being a token of some SLE may be an extrinsic property of an object; (3) it is reasonable to expect the set of incarnations of a SLE to be gerrymandered. (4) The task is to explain what makes an entity a token of a SLE, and what makes an acoustic blast an incarnation of a SLE. And (5), this involves giving the generic incarnation conditions.

Though (GIC) is often not clearly formulated, it is mainly this question to which a variety of accounts have been given. These include convention-based accounts, intention-based accounts, response dependent-based accounts, functional accounts, and others.

#### **4. The distinction between grounds and conditions**

I now want to propose a distinction between two sorts of factors that figure into determining whether an entity is a token of a type. The distinction is between what the *conditions* are for tokening, and what *grounds* these tokening conditions. I will further distinguish different aspects of the tokening conditions. The full breakdown of determining factors I will discuss is depicted in figure 2:

Fig. 2



With the conditions/grounds distinction (and some subsidiary distinctions) we can clarify what is mistaken about certain approaches to the tokening question, and can approach the question in a more promising way. In particular, it helps to distinguish the roles that different sorts of factors external to the incarnations play in determining type-membership.

I will discuss this distinction in connection with taking a convention-based approach to the tokening question, as Michael Devitt and Herman Cappelen do. Convention turns out to be a useful case for clarifying the distinction between grounds and conditions. Conversely, making the distinction clear shows that convention is not a plausible explanation for providing the generic tokening conditions for words.

#### 4.1 *The intuitive distinction*

Consider for a moment written words, rather than uttered tokens. What makes an ink mark printed on a piece of paper a token of some particular word? The traditional type/token model argues that it is strictly a matter of the intrinsic properties of the mark.<sup>25</sup> A mark is a token of ‘Aristotle’ if it has the proper shape, however that shape was produced. A different approach is taken by David Kaplan among others, who argues that the intention of the person making the mark is the central requirement for a mark counting as a

---

<sup>25</sup> Cf. Cappelen 1999.

token of 'Aristotle'.<sup>26</sup> He argues that it is absurd that a mark should count as a token of 'Aristotle' if it is made by a wave washing over a beach. Instead, tokening 'Aristotle' is a matter of there being the right intention in place, in the marking. This has the virtue of showing how illegible or nonstandard marks may nonetheless count as tokens of 'Aristotle'. Herman Cappelen has recently argued against Kaplan's view, countering that intention is neither necessary nor sufficient for tokening. One argument he gives, for instance, is that if producer-intention is a condition for a mark to be a word, then a sentence stating a semantic fact such as "'Alice' refers to Alice" is necessarily uninformative, which counterintuitively would make knowing such facts useless for knowledge of a language.<sup>27</sup>

It is not my aim here to consider the merits of these views, but to make use of them for classifying different ways we might answer the tokening question. The non-intentional view gives a straightforward set of conditions for a mark to be an incarnation of a given word: the conditions are strictly physical. All that is required of a mark is that it have the right intrinsic physical characteristics. Now, suppose someone were to ask a non-intentionalist, in virtue of what are these the conditions? What makes it the case that a mark having such-and-such physical characteristics counts as a token of a given word? One reasonable answer that might be given is that the conditions are grounded in *convention*. This is the view that Cappelen, for instance, puts forward, and he criticizes the traditional type/token model for neglecting the question:

*It is a matter of convention that what you look at right now are tokens of words. They would not be word tokens if it hadn't been for the presence of certain kinds of conventions. Such conventions are upheld by*

---

<sup>26</sup> Kaplan 1990. Millikan 1984, Bromberger 1989, Richard 1990, Barber 2006 also in different ways advocate intention to be a key condition for tokening. There are many variants on these views. Kaplan's is the most clearly focused on intention alone. There are other views that are more explicitly hybrids, such as Alward 2005. Millikan's, Richard's, and Barber's views may also properly be understood as hybrids.

<sup>27</sup> Cappelen 1999, pp. 97, 101.

*intentional linguistic activity. So, there can be word tokens only if there is intentional linguistic activity. It's a mistake to infer intentionalism [i.e., Kaplan's view] from this. The conventions we have are, very roughly, of two kinds:*

*(C1) Entities with such-and-such properties count as tokens of the same sign*

*(C2) Entities with properties P (where entities with P count as tokens of the same sign according to some C1-type convention) count as tokens of the same sign as tokens with properties P' (where P' entities count as tokens of the same sign according to some C1-type convention).*

*C1-type conventions make it the case that these two tokens are tokens of the same-sign: **lobster lobster**.*<sup>28</sup>

For our purposes, the C1-type conventions are the pertinent ones. (In speaking of C2-type conventions, Cappelen is concerned to explain why tokens within different sign systems, such as spoken, written, Braille, semaphore, etc., all count as tokens of the same sign.)

Below, I will suggest that Devitt holds along with Cappelen that tokening conditions are grounded in C1-type conventions, although they differ slightly on what the tokening conditions are. And I will argue that grounding in C1-type convention is implausible, so both Devitt and Cappelen are mistaken. But first let me highlight a distinction Cappelen implicitly uses.

Notice that the role for C1-type conventions is not restricted to the non-intentionalist view. As Cappelen goes on to say:

*The disagreement between intentionalists and non-intentionalists is what to put in for "such-and-such" in "such-and-such properties" in C1. The intentionalist says this should include reference to the intentional production history. The non-intentionalist denies this. Both views are compatible with the obvious fact that concrete particulars are word tokens only because of the presence of such conventions and the equally*

---

<sup>28</sup> Cappelen 1999, p. 99.

*obvious fact that such conventions require intentional activity on the part of sign users.*<sup>29</sup>

As Cappelen points out, convention might play the same role in the intentionalist camp as it does in the non-intentionalist camp. Like the non-intentionalist, the intentionalist could hold that we establish a convention in our linguistic community, and that that convention is that any mark produced with those intentions counts as an incarnation of a given word. On some versions of intentionalism, this might be a strange response, since presumably the intention-theorist would have some independent reason to take intention to be a condition (or the sole condition) for a mark to count as an incarnation of a given word. But it is not impossible to hold this view. An intentionalist might, for instance, hold that there are a number of alternative conditions that could determine which marks count as incarnations of a given word, but we happen to have a convention that the conditions include intentional ones.

As I mentioned, I will argue below that Cappelen is wrong about the “obvious fact” that C1-type conventions establish the conditions for tokening. He is not wrong, however, to distinguish (a) the “such-and-such properties” that a mark needs to satisfy to count as an incarnation of a given word from (b) the factors that ground or determine the fact that these “such-and-such properties” are the conditions that a mark needs to satisfy. I will call the “such-and-such properties” the *tokening conditions*, and the factors that ground those tokening conditions I will call the *grounds*. In the case of convention, there is a clear difference between the work we do in giving the conditions for a mark to be a token of a word and the work we do in giving the conditions for the particular convention to be in place that makes it the case that a mark satisfying those conditions counts as tokening the word. Cashing out the intuitive distinction is a difficult matter, and warrants further investigation. But for the present purposes, the intuitive distinction already lets us begin to clarify different approaches to the tokening problem.

---

<sup>29</sup> Cappelen 1999, p. 99.

Let us consider where Devitt's view fits in. In introducing his convention-based account of word tokening, Devitt is fond of using the example of a yellow flag signaling the presence of yellow fever on ships. He gives a convention-based account of *signaling yellow fever*: the role convention plays is to set up the fact that the satisfaction of certain conditions counts as a signal of yellow fever. Just this much, however, does not make it clear what Devitt takes the conditions to be for being a signal of yellow fever. The conditions themselves may be strictly intrinsic to the token – they may involve simply that a yellow flag be waving on the mast of a ship. Or they may be extrinsic. If the convention is set up so that some social or intentional facts be in place in order for the flag to count as signaling yellow fever, then there are some extrinsic tokening conditions as well intrinsic ones.

I will try to pin down Devitt's position just below. But even while the tokening conditions he endorses are undetermined, we can already see that it is helpful to distinguish grounds from tokening conditions in classifying views:

Table 1:

	<b>Grounds</b>	<b>Tokening conditions</b>
<b>Type/token</b>	?	Strictly intrinsic
<b>Cappelen</b>	Convention	Strictly intrinsic
<b>Cappelen's characterization of intentionalists</b>	Convention	Strictly intentional
<b>Devitt</b>	Convention	?
<b>Kaplan's actual view<sup>30</sup></b>	?	Mostly intentional

<sup>30</sup> As Alward 2005 points out, Cappelen's characterization of Kaplan is too strong. Kaplan may be reasonably interpreted as holding that the tokening conditions are largely but not entirely intentional. (Kaplan has confirmed this in personal communication.)

As Cappelen points out, at least two different sorts of tokening conditions may be grounded in convention. But equally, when we attempt to classify more proposals below, we will see that there may be different grounds for one particular set of tokening conditions. One could be an advocate of conventional grounding without yet taking a position on the tokening conditions, or conversely, advocate a particular set of tokening conditions without taking a position on the grounding for those conditions.

#### 4.2 *Conventional grounds and tokening conditions*

In distinguishing grounds from tokening conditions, I have suggested that the tokening conditions may be intrinsic even if the grounds are extrinsic. On Cappelen's view, for instance, tokening conditions are strictly intrinsic while those conditions are grounded in convention. On a reasonably standard understanding of convention, however, this is not quite right. Just as many people take it to be absurd that an **Aristotle**-shaped mark made by a wave in the sand would count as a token of 'Aristotle', many also will take it as absurd that a yellow flag hanging from a mast should count as participating in a convention, if it is not hung with the intention of participating in the convention. This implies that a conventional grounding may require at least some extrinsic tokening conditions.

Extrinsic tokening conditions, however, should not all be dropped into one bucket. One further distinction is needed. Even having separated out the grounds from the tokening conditions, there are two different kinds of tokening conditions. This is again particularly clear for tokening a convention-grounded type: (i) there are the conditions that are generically required to participate in a convention, and (ii) there are the conditions that are set out by the particular convention, which tokens need to satisfy.<sup>31</sup> For short, we will say that the tokening conditions break down into *participation* conditions and *type-specific* conditions.

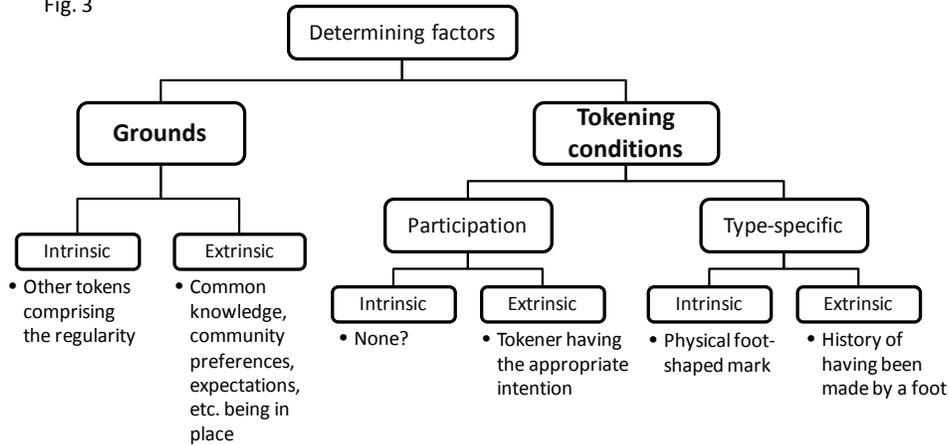
---

<sup>31</sup> The distinction can be made more general than just for application to convention-based grounding, but limiting it to convention is clearer for present purposes.

Consider some examples. Let us suppose that there is a convention in place that a place-setting of silverware laid out in a particular way (e.g., fork on the left, knife and spoon on the right), counts as being correctly laid out. Let us further suppose that there is an intentional requirement for participating in a convention. I.e., to count as being correctly laid out, a place setting must have been laid out with the intention to conform to the convention. But suppose that those are the only extrinsic conditions: apart from the requirements for participation, the convention itself sets out only intrinsic conditions for counting as being correctly laid out.

This can be contrasted with cases in which there are both intrinsic and extrinsic conditions set out by the convention. Suppose that we have a convention that if a person leaves a footprint in a certain place, it counts as a territorial claim. Now, the property *being a footprint* is an extrinsic property of a physical mark: it requires that it have been put in place by a foot in the past. What, then, are the conditions for *being a territorial claim*? First there are the conditions, internal and external to the mark itself, that determine that the mark participates in the convention. These are the ones that involve the tokeners' intention. The participation factors are those that are generically required for instancing a convention, regardless of what the convention is. Then there are the type-specific conditions set out by the convention. In contrast to the place-setting case, the conditions for a mark to count as a territorial claim are both internal and external to the mark itself. Namely, those that determine that the mark is a footprint in the right place. Filling out all the conditions, the factors break down into:

Fig. 3



In short, even when the participation conditions are extrinsic, the type-specific conditions may be strictly intrinsic, or they may be strictly extrinsic, or they may be both intrinsic and extrinsic. It is usually when there are extrinsic type-specific conditions, that we take a type itself to be relational. But if the type-specific conditions are strictly intrinsic, we often do not consider the type to be relational, even if there are extrinsic participation conditions.

Notice that when Cappelen talks about “such-and-such properties” in C1-type conventions, these include all the tokening conditions, participation and type-specific. Cappelen argues that once the relevant conventions are in place, a wave washing over the beach in the shape **Aristotle** is a token of ‘Aristotle’. This means that he takes there to be no participation conditions for tokening, but only intrinsic type-specific conditions.

Devitt disagrees.<sup>32</sup> He holds that there are intentional participation conditions. The most plausible interpretation I can give of Devitt’s view is that he takes the yellow flag case to be like the place-setting case. That is, he takes the grounds to be conventional, and he also takes there to be intentional participation conditions for a tokening, but takes those to be the only intentional conditions. That is, he takes the extrinsic factors involved to be

---

<sup>32</sup> Personal communication.

(1) the extrinsic factors involved in determining that the convention is in place, and (2) the extrinsic factors involved in a particular token counting as participating in that convention. Apart from these, he takes the conditions for an object counting as signaling yellow fever to be strictly intrinsic. So Devitt's view of yellow flag signals seems to be very much like Cappelen's except that he requires of a token that it have been made with an intention to participate.

This also means that the property *being a territorial claim* is relational in a sense that *being a signal of yellow fever* is not. The source of relationality for signaling yellow fever is only in the participation conditions, while the sources of relationality for *being a territorial claim* involves type-specific extrinsic conditions such as whether the mark was made by a foot, in addition to the participation conditions.

Although this is probably a reasonable interpretation of the yellow flag example, it is not clear how Devitt intends to apply it to SLEs, and hence what he takes to be the source of their putative relationality. Are SLEs relational in virtue of their being like yellow flags? If so, then the only external conditions for an acoustic blast to count as an incarnation of a word are those involved in participating in a convention. That would mean that the convention sets out strictly intrinsic type-specific conditions for the blast to satisfy. Or are they relational in virtue of having to satisfy further external conditions? And if so, then what are those conditions?

#### 4.3 *Classifying views*

Having made these distinctions, it is possible to classify views on tokening conditions for SLEs in a much more useful way than just between the types of intrinsic and extrinsic conditions they take to be involved. Table 2 is a first pass at categorizing the factors that different views propose. Each view is broken down according to how it would fill out figure 2, as applied to the generic tokening of SLEs. By "intrinsic" factors I mean those that are intrinsic to the incarnations of tokens of the type in question, and "extrinsic" factors are those that are not intrinsic to the incarnations.

Table 2:

	<b>Grounds</b>	<b>Tokening conditions: Participation</b>	<b>Tokening conditions: Type-specific</b>
<b>Traditional type/token</b>	?	None	<b>Intrinsic:</b> Acoustic <b>Extrinsic:</b> None
<b>Kaplan</b>	?	None?	<b>Intrinsic:</b> None <b>Extrinsic:</b> Intention to reproduce, historical conditions
<b>Cappelen</b>	Convention	None	<b>Intrinsic:</b> Physical conditions <b>Extrinsic:</b> None
<b>Devitt</b>	Convention	<b>Intrinsic:</b> ? <b>Extrinsic:</b> Intention to produce meaningful token?	<b>Intrinsic:</b> Physical conditions <b>Extrinsic:</b> ?
<b>Alward</b>	?	None	<b>Intrinsic:</b> Acoustic (that change diachronically) <b>Extrinsic:</b> Intention to produce a token; competence of speaker
<b>Barber</b>	Nature of word-use as an illocutionary act	N/A	<b>Intrinsic:</b> intrinsic features for satisfying constraints on being judged to have tokened a word <b>Extrinsic:</b> Illocutionary intention; extrinsic features for satisfying constraints on being judged etc.
<b>Hearer-dependence conditions (Miscevic)</b>	?	N/A?	<b>Intrinsic:</b> None <b>Extrinsic:</b> Normal hearer disposed to interpret as K
<b>Hearer-dependence grounding</b>	Normal hearer disposed to interpret as K; the structure of the normal hearer	N/A?	<b>Intrinsic:</b> Physical conditions <b>Extrinsic:</b> None
<b>Causal role functional</b>	?	?	<b>Intrinsic:</b> None <b>Extrinsic:</b> Satisfy causal role conditions in appropriate context
<b>Teleofunctional</b>	History of reproductively established family (REF)	<b>Intrinsic:</b> ? <b>Extrinsic:</b> Appropriate connection with REF	<b>Intrinsic:</b> None <b>Extrinsic:</b> Intention to reproduce
<b>Davidsonian</b>	Being an interpreter of a language	N/A	<b>Intrinsic:</b> None <b>Extrinsic:</b> Intention
<b>Etc.</b>			

Putting views into this table is somewhat dodgy, at least in part because it requires precisifying conditions that the views themselves leave ambiguous. Nonetheless, it helps show both where the claims of a given view are directed, and where they may need to be supplemented.

## 5. Devitt and Cappelen on convention and SLEs

Like Cappelen, Devitt argues that the tokening conditions for SLEs are grounded in convention. In a recent reply to critics, he writes:

*Sometimes relational properties are correlated well with superficial properties and hence their presence is easily detected, but sometimes they are not. Thus it is fairly easy to detect money, cars and echidnas, but not so easy to detect Australians, the unemployed, smokers, Chomskians, paperweights and moons. And it is mostly easy to detect SLEs because they are established by conventions that correlate linguistic properties with superficial ones. The clues to a linguistic property are clues to it because conventions bestow that property on objects that provide those clues.*<sup>33</sup>

In the same context, he again clarifies the point with the “yellow fever” example:

*Indeed, think of the conventions for conveying complex messages with a simple symbol like a flag: a yellow flag on a ship’s mast, meaning **This ship has yellow fever**, has the property of referring to yellow fever even though there is no one spot on the flag that does so. We can create conventions that make structure as implicit or explicit as practicality dictates.*

I will criticize the view that the conditions for being a token of a SLE can in fact be grounded in convention, using effectively the same argument that Strawson used against Austin, in “Intention and Convention in Speech Acts” of 1964.

---

<sup>33</sup> Devitt 2006a

### 5.1 Strawson's critique of convention

Strawson's argument against Austin is well known, so I will just give a quick summary. In *How to Do Things with Words*, Austin argued that speech acts are conventional. He used a couple of examples – acts of christening, such as christening a boat by using a bottle of champagne, and acts of marrying – as the key prototypes for how speech acts work. These, he argued, involve conformance to convention. Strawson replied that while the convention interpretation may be correct for those particular examples, they should not be understood as the prototypes of speech acts. Most speech acts do not involve convention, on Strawson's view.

What, Strawson asks, makes it that case that “don't go!” is an entreaty? He points out that we do not need a convention to be established in order to ground the illocutionary force of “don't go!” What make it an entreaty are factors relating to my situation, my attitude, to the person I am saying it to, my manner, and my current intentions. Strawson argues:

*To suppose that there is always and necessarily a convention conformed to would be like supposing that there could be no love affairs which did not proceed on lines laid down in the Roman de la Rose, or that every dispute between men must follow the pattern specified in Touchstone's speech about the countercheck quarrelsome and the lie direct.<sup>34</sup>*

On Strawson's view, we do not need convention for most speech acts. When do we need grounding in a convention? Only when we cannot derive the same outcome on more general grounds.

There are really two Strawsonian roots in that paper for criticizing convention. One of them is to say, is convention plausibly the explanation? For instance, would it require such a detailed set of conventions that it is implausible that we could even have established the required conventions? And the other is to provide an alternative explanation, to propose that the illocutionary force of an utterance can be grounded with things like intentions and the situational facts and so on. As he puts it:

---

<sup>34</sup> Strawson 1964

*I would say that the illocutionary force of the speech act is not grounded in convention, since it does not need convention to give it the force it has. No knowledge, agreement, etc.*<sup>35</sup>

Strawson's proposal, filled out in detail particularly by Bach and Harnish,<sup>36</sup> is that speech acts are grounded along Gricean lines, i.e., in the communicative intentions of the speakers.

## 5.2 Applying the critique to SLEs

Now, it looks as though Strawson's critique will not apply to standard linguistic entities. The reason it looks this way is that there plausibly are some conventions involved in how a word like 'Aristotle' is typically tokened. It is an arbitrary choice that assigns an **Aristotle**<sub>1</sub>-like sound pattern to the person Aristotle, and this association may plausibly be understood as having instituted a convention.<sup>37</sup>

However, this is not the relevant point, when it comes to answering the tokening question. Recall the hierarchy of types and the generic tokening question. The question we are asking is not: What are the conditions for an acoustic blast to be an incarnation of 'Aristotle'? Rather, the relevant question is: Given a word W, what are the conditions for an acoustic blast to be an incarnation of W?

To answer this question, we want to find a generic set of conditions for an acoustic blast to be an incarnation of a given word. We expect that it will be the same principle for **cat**<sub>1</sub> and **cat**<sub>2</sub> to be tokens of 'cat' as for **dog**<sub>1</sub> and **dog**<sub>2</sub> to be tokens of 'dog' and **Aristotle**<sub>1</sub> and **Aristotle**<sub>2</sub> to be tokens of 'Aristotle'. Now, even if there is a convention involved in fixing the word 'Aristotle', that is not relevant to the generic conditions. Instead, I want to suggest that

---

<sup>35</sup> Strawson 1964, pp. 444-445.

<sup>36</sup> Bach and Harnish 1979

<sup>37</sup> This is not the only, or even the prevailing, interpretation of what takes place in a reference-fixing event. But as I point out here, this question is orthogonal to an answer to the generic tokening question.

the conventional grounding for the relevant question – i.e., the generic tokening question – is extremely implausible on exactly the same grounds that Strawson gives.

To simplify, let us suppose that when we introduce the word ‘cat’ or ‘dog’ or ‘Aristotle’, we associate the referent with an acoustic prototype. Let us further assume that tokening involves satisfying (i) an intentional participation condition (which may include a historical chain), and (ii) being similar in some statistical way to the respective prototype. Those two conditions are all the membership conditions, so for a particular token those are what it takes to determine whether an acoustic blast is or is not a token of the word ‘cat’ or ‘dog’ or ‘Aristotle’. This probably is not the right account, but it is an account of tokening.

Consider the statistical similarity metric. This metric (or a function on the metric) will have to be quite complicated, in order to give any reasonable results at all for what counts as an acceptable token. There will, for instance, have to be different tolerances depending on whether the speaker is native, or whether the speaker is a child, or whether the speaker has an accent, and so on. An enormous number of factors will have to figure into the metric. These make it implausible that we could ground the tokening conditions in convention even if we wanted to. Moreover, even the idea that the type-specific incarnation conditions are just a matter of statistical proximity to a prototype is surely a gross simplification. The establishment of convention for deriving the generic tokening conditions is implausibly difficult.

Moreover, the other Strawsonian point is even more relevant: conventional grounding is unnecessary. What makes it the case that a slurred token uttered by a drunk person and a babyish token uttered by a child and a creaky token uttered by a nonagenarian are all three tokens of ‘cat’, which also makes it the case that three different utterances by these people are all tokens of ‘dog’? The explanation might mostly have to do with the intention of the speaker to generate the token, or the ability of the hearer to interpret it, or with the physical characteristics of the larynx, mouth, air, ear, and brain, or perhaps with all of these. But there is no need to introduce convention in

addition to these factors, any more than convention needs to be added to make it the case that “don’t go!” counts as an entreaty. The gerrymandered similarity metric may plausibly be grounded in any or all of the above factors, without having to invoke convention.

Cappelen says that it is an “obvious fact that concrete particulars are word tokens only because of the presence of such [*i.e. C1-type and C2-type*] conventions.” The problem is that Cappelen fails to distinguish the generic incarnation conditions from the question of the incarnation conditions for a particular word such as ‘cat’. It is plausible that a convention is established in the fixing of the word ‘cat’, for instance connecting it to an acoustic prototype. But that is the least part in the explanation of what the conditions are for a particular acoustic blast **cat**<sub>1</sub> to count as a token of ‘cat’. Or as Cappelen puts it, for determining what the same-sign relation is. It is not only non-obvious, but surely false, that if there is a ground for the same-sign relation which applies generically across words, that it is grounded in convention. To explain why **cat**<sub>1</sub> and **cat**<sub>2</sub> are tokens of ‘cat’ and why **dog**<sub>1</sub> and **dog**<sub>2</sub> are tokens of ‘dog’ and why **Aristotle**<sub>1</sub> and **Aristotle**<sub>2</sub> are tokens of ‘Aristotle’, the grounding is surely non-conventional.

Most of the work of this paper has been to lay the groundwork for understanding what the basic requirements are for an account of SLE-tokening. I have not addressed the virtues and shortcomings of most of the accounts I listed in table 2. From what I have argued here, any of them remains eligible to provide at least part of an account of the grounds or tokening conditions. Except convention.

## REFERENCES

- Alward, P., Between the Lines of Age: Reflections on the Metaphysics of Words, *Pacific Philosophical Quarterly*, 86 (2005): 172.
- Bach, K., and R. M. Harnish, *Linguistic communication and speech acts*. Cambridge, Mass.: MIT Press, 1979.
- Barber, A., Testimony and Illusion, *Croatian Journal of Philosophy*, 6:18 (2006): 401.
- , Sentence Realization Again: Reply to Rey, *Croatian Journal of Philosophy*:23 (2008): 233.
- Bromberger, S. 'Types and Tokens in Linguistics.' Reflections on Chomsky. Ed. George, A. Oxford: Blackwell, 1989. 58.
- Bromberger, S., and M. Halle. 'The Ontology of Phonology (Revised).' Phonological Knowledge: Conceptual and Empirical Issues. Eds. Burton-Roberts, N., P. Carr and G. Docherty. Oxford: Oxford University Press, 2000. 19.
- Cappelen, H., Intentions in Words, *Nous*, 33:1 (1999): 92.
- Collins, J., A Note on Conventions and Unvoiced Syntax, *Croatian Journal of Philosophy*:23 (2008): 241.
- Devitt, M., Defending *Ignorance of Language*, *Croatian Journal of Philosophy*, 6 (2006a): 573.
- , *Ignorance of Language*. Oxford: Clarendon Press, 2006b.
- Devitt, M., and K. Sterelny, *Language and Reality*, 2nd ed., Oxford: Blackwell, 1999.

Fine, K., The Nonidentity of a Material Thing and its Matter, *Mind*, 112 (2003): 195.

Gibbard, A., Contingent Identity, *Journal of Philosophical Logic*, 4 (1975): 187.

Haslanger, S., Ontology and Social Construction, *Philosophical Topics*, 23:2 (1995): 95.

Kaplan, D., Words, *Proceedings of the Aristotelian Society*, 64 (1990): 93.

Liberman, P., *Speech: A Special Code*. Cambridge: MIT, 1996.

Millikan, R. G., *Language, Thought, and Other Biological Categories: New Foundations for Realism*. Cambridge: MIT Press, 1984.

Rea, M. 'Material Constitution.' Lanham, MD: Rowman & Littlefield, 1997.

Rey, G., Conventions, Intuitions and Linguistic Inexistence: A Reply to Devitt, *Croatian Journal of Philosophy*, VI (2006a).

---. 'Empty Representations in Linguistic Perception.' *Early Content*. 2006b. 1.

---. 'The Intentional Inexistence of Language -- But Not Cars.' *Contemporary Debates in Cognitive Science*. Ed. Stainton, R. Oxford: Blackwell Publishers, 2006c. 237.

---. 'Externalism and Inexistence in Early Content.' *Prospects for Meaning*. Ed. Schantz, R. New York: deGruyter, 2007.

---, In Defense of Folieism: Replies to Critics, *Croatian Journal of Philosophy*:23 (2008): 177.

Richard, M., *Propositional Attitudes*. Cambridge: Cambridge University Press, 1990.

Strawson, P. F., Intention and Convention in Speech Acts, *Philosophical Review*, 73:4 (1964): 439.

Szabo, Z. G., Expressions and their Representations, *Philosophical Quarterly*, 49:195 (1999): 145.