

[Penultimate draft, final version forthcoming in *Inquiry*]

Shareability of Thought and Frege's Constraint: A Reply to Onofri

Romain Bourdoncle 

Institut Jean Nicod, Département d'études cognitives, École

Normale Supérieure

`romain.bourdoncle@ens.fr`

Abstract

Onofri (2018) proposes an individuation criterion for thoughts that purports to satisfy both shareability (the notion that different thinkers, or a single thinker at different times, can and generally do think type-identical thoughts) and Frege's constraint (according to which two thoughts are different if it is possible for a rational subject to endorse one while rejecting the other). I argue that his proposal fails to satisfy Frege's constraint. Then I propose a modification to Onofri's proposal to fix the problem.

Keywords: Content · Frege's Constraint · Communication · Frege cases · Coreference · Relationism

1 Onofri's Criterion

We would like thoughts to be shareable (i.e. such that they can be, and often are, type-identical across agents) and we would like them to play a role in rationalizing psychological explanation. For the latter, any two thoughts such that it is possible for a rational subject to endorse one while rejecting the other should be counted as different. This is, roughly, Frege's constraint ((FC) hereafter). A consequence of (FC) is that the *same-thought-as* relation must be finer than referential equivalence. On the other hand, for thoughts to be shareable, the *same-thought-as* relation must yield a partition properly coarser than the partition into singletons. Thus (FC) and shareability pull the granularity of thought individuation in opposite directions. As a result, it is not straightforward how to individuate thoughts so that they are both shareable and fine-grained enough to satisfy (FC). In particular, one may wonder: Does shareability require an individuation criterion that is properly coarser than the one required to satisfy (FC)? If the answer is yes, then it would turn out that shareability is not compatible with (FC).

Onofri (2018) proposes an individuation criterion for thoughts that purports to satisfy both constraints. His proposal is in three steps. First, Onofri defines what he calls the 'linking relation', as follows¹:

Linking Relation (*L*) Two thoughts t_a and t_b stand in *L* iff the thinker of t_a and the thinker of t_b know that t_a and t_b ascribe the same property to

¹Onofri's notion of linking is similar to Recanati (2012)'s except that the latter is restricted to the intrasubjective domain.

the same object.

Then Onofri provides a first-pass individuation criterion in terms of (L) , as follows:

(IC, first pass) A thought t_a is the same thought as a thought t_b iff t_a and t_b stand in L .

As it turns out, L is not a transitive relation. To show this, here is an example.² Consider three thinkers **A**, **B** and **C**. **B** knows that Superman = Clark and **B** knows that Superman = Kal-El. **A** only knows that Superman = Clark. **C** only knows that Superman = Kal-El. In a context in which the identity of Clark and Superman is common ground, **A** tells **B** 'Clark can fly'. Call t_A , **A**'s thought on that occasion and t_{B_1} , **B**'s thought on that occasion. Then t_A and t_{B_1} stand in L .³ In a context in which the identity of Kal-El and Superman is common ground, **C** tells **B** 'Kal-El can fly'. Call t_{B_2} , **B**'s thought on that occasion, and t_C , **C**'s thought on that occasion. Then t_{B_2} and t_C stand in L . However, it does not follow that t_A and t_C stand in L .

Onofri does not offer a clear diagnosis of why transitivity fails. The failure of transitivity in this case may be unpacked as follows: **A** does not know

²I am reusing Onofri's original example, except that I substitute identity thoughts (and utterances of the form ' a is b ') for predicative thoughts (and utterances of the form ' a is F '). Since L is formulated in terms of predicative thoughts, I think the example is more straightforward put this way.

³As far as I understand the criterion, that t_A and t_{B_1} stand in L is to be unpacked as follows: **A** and **B** both know that the thought that **A** expresses has the same referential content than the thought that **B** entertains when **B** understands **A**'s utterance. I will assume this reading pattern of L in what follows. (For stylistic reasons, I will keep this reading implicit most of the time from now on). Moreover, in this example, the content is of course pseudo-referential. I ignore this complication, as Onofri does, for present purposes.

that the thought she expresses by 'Clark can fly' corefers with the thought **C** expresses by 'Kal-El can fly' *or* **C** does not know that the thought she expresses by 'Kal-El can fly' corefers with the thought **A** expresses by 'Clark can fly'. There are various possible reasons why **A** or **C** may fail to know that their respective thoughts t_A and t_C corefer, which are of varying significance for Onofri's criterion. Let me mention two.

An obvious sufficient reason for transitivity to fail in this case is that **A** and **C** may not have been present during each other's utterances to **B**. For all that the example says, **A** may be ignorant of the existence of t_C or **C** may be ignorant of the existence of t_A . In effect, this seems to be a sufficient reason for them not to believe (hence not to know) that the thoughts they express by their respective utterances corefer.

Note the consequence of this for Onofri's criterion. Onofri's criterion is supposed to be a necessary and sufficient condition for two thoughts to be the same. In particular, it is supposed to be a necessary condition for two thoughts to be the same. Hence, if you are in Paris in 1886 and I am in Paris in 2020 and we both innerly assert ' $2 + 2 = 4$ ', we fail to think the same thought, on this criterion. This is because we don't believe (a fortiori do not know) anything of each other's occurrent thoughts. But it seems that type-identity between thoughts should not be thus contingent on e.g. which conversations we have. To remedy this on the proposed criterion, one could

try to define L in dispositional terms instead.⁴ Alternatively, one could reformulate the criterion so that it is not intended as a necessary condition for sameness of thought in general.

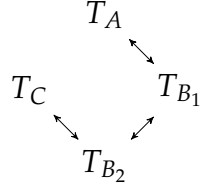
Another sufficient reason for transitivity to fail here is that **A** does not know that 'Clark' and 'Kal-El' corefer. As a result, **A** may fail to know that the thought **C** expresses by 'Kal-El can fly' corefers with the thought **A** expresses by 'Clark can fly'. Similar considerations apply to **C**.

Be that as it may, L is not transitive, but identity is transitive, therefore L is not a candidate for the *same-thought-as* relation as it stands. Hence the third step: to remedy this, Onofri considers the transitive closure of L – which he calls 'the indirect linking relation' (L^*). He then proposes to redefine **(IC)** in terms of L^* , as follows:

(IC) A thought t_a is the same thought as a thought t_b iff t_a and t_b stand in L^* .

Here is a graph to illustrate how L^* is supposed to help with the failure of transitivity. The graph relation is L , and L^* is connectedness on the graph.

⁴As (Cappelen and Hawthorne, 2009, 60) remark, there is a sense of agreement and disagreement that applies to 'interaction-free pairs of individuals so long as there is some view about the world that they share.' A non-interactive notion of sameness of thought is arguably needed to define a notion of agreement and disagreement in this sense.



Node T_A and node T_C are not related on the graph, but they are connected (i.e. reachable from each other). This reflects the fact that, although **A** and **C**'s thoughts do not stand in the direct linking relation (L), they stand in the indirect linking relation (L^*). Hence transitivity is restored by L^* .

In the next section, I argue that **(IC)** in terms of L^* is too coarse to satisfy Frege's constraint.

2 (IC) Is Too Coarse To Satisfy Frege's Constraint

For convenience, I will re-use Kripke's well-known Peter example.⁵ Also for convenience I rebaptize Peter '**P**'. **P** (aka Kripke's Peter) believes that London is not pretty and he also believes of London (under the French name 'Londres'), that it is pretty. By hypothesis, **P** is rational.

Now consider two other protagonists: there is **Q**, who, like **P**, is bilingual in French and English. **Q** knows that 'London' and 'Londres' corefer. There

⁵Kripke (1979).

is also **R**, a normal monolingual English speaker competent with the name 'London'.

P tells **Q** 'Londres est jolie'. Call t_{P_1} and t_{Q_1} , **P** and **Q**'s thoughts on that occasion, respectively. They are linked. At some other time, **Q** tells **R** 'London is pretty'. Call t_{Q_2} and t_{R_1} , **Q** and **R**'s thoughts on that occasion, respectively. They are linked. Note that t_{Q_1} and t_{Q_2} also are linked, by hypothesis.⁶ At some other time, **R** tells **P** 'London is pretty'. Call t_{R_2} and t_{P_2} , **R** and **P**'s thoughts on that occasion, respectively. They are linked. (Note that t_{R_1} and t_{R_2} also are linked). Of course, **P** disagrees with **R**, for he disbelieves of London, under the name 'London', that it is pretty.

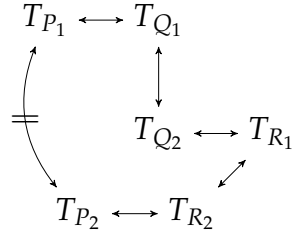
P can rationally reject the thought he associates with the utterance 'London is pretty' (i.e. t_{P_2}) while endorsing the thought he associates with the utterance 'Londres est jolie' (i.e. t_{P_1}) because **P** does not know that t_{P_1} and t_{P_2} corefer. In other words, t_{P_1} and t_{P_2} are *unlinked* from **P**'s perspective. Now recall what **(FC)** says:

(FC) Two thoughts are different if it is possible for a rational subject to endorse one while rejecting the other.

By **(FC)**, t_{P_1} and t_{P_2} are different. By **(IC)**, t_{P_1} and t_{P_2} are the same. I conclude that **(IC)** violates **(FC)**, because **(IC)** is too coarse.

⁶I assume that the respective memories of **P**, **Q** and **R** work properly. I also assume that the protagonists are lexically competent with the adjective 'pretty', etc. throughout the episode.

Here is another graph to illustrate how **(IC)** and **(FC)** clash with each other in the present case. The graph relation is L , L^* is connectedness on the graph, and crossed out edges are used to stress disconnectedness on the graph.



On the one hand, node T_{P_1} and node T_{P_2} are connected by a path of more than one edge on the graph. This reflects the fact that t_{P_1} and t_{P_2} stand in the indirect linking relation (L^*). On the other hand, node T_{P_1} and node T_{P_2} are disconnected on the graph, as stressed by the crossed out edge between them. This reflects the fact that t_{P_1} and t_{P_2} are *unlinked* for **P**.

3 A Diagnosis & Recipe for Constructing Similar Cases

Intuitively, when things go well, L^* purports to be a route by which a thought is transmitted from one speaker to another.⁷ But in general, we have no reason to assume that L^* -continuity will be consistent with the perspective each of the thinkers has on the thoughts they respectively deploy

⁷Hence the similarity with Kripke's notion of a causal-historical chain (Kripke (1980)), as Onofri himself rightly notes.

along a chain. For how thoughts are linked and unlinked for a thinker may be idiosyncratic, and may change over time.

By extending the linking relation (L) via communication chains, one thus naturally exposes the individuation criterion based on it to Frege cases. This is another instance of the familiar observation that shareability and **(FC)** pull the granularity of thought individuation in opposite directions. Hence, the described counterexample should come as no surprise. If one wants to individuate thoughts by membership in the ordered sets corresponding to L^* -routes, one will violate **(FC)** as soon as a route includes two thoughts of a single thinker that are unlinked from the perspective of the thinker to which they belong.

4 An Attempt to Fix the Problem

If the diagnosis offered in the previous section is on the right lines, then a way to solve the problem suggests itself.

I repeat the diagnosis: **(IC)** violates **(FC)** whenever two coreferential thoughts are deployed by a single individual along a L^* -route, if these thoughts are unlinked from the perspective of that individual. Therefore, to respect **(FC)**, we want a L^* -route that is compatible, instead, with how thoughts are linked and unlinked from the perspective of their thinkers.

In other words, if we want to satisfy **(FC)**, we do not want to link thoughts belonging to one individual on a L^* -route if these thoughts are unlinked for that individual. For to do so goes against **(FC)**. Instead, to respect **(FC)**, I suggest that we should consider as linked as many thoughts as possible in the way of L^* , while refusing to link thoughts of a single individual on a L^* -route if these thoughts are unlinked from the perspective of that individual.

Here is a version of the indirect linking criterion that incorporates the point just made:

(Indirect Linking modulo FC) Two thoughts t_a and t_b stand in the indirect linking relation iff there is a tuple $\langle t_a, \dots, t_n, t_b \rangle$ such that:

- (i) each member stands in L to its successor;
- (ii) no thoughts of a single individual that are unlinked for that individual belong to $\langle t_a, \dots, t_n, t_b \rangle$.

This redefinition of the indirect linking relation looks stipulative as it stands. This is because **(IC)** together with clause (ii) essentially says 'count the thoughts as the same unless there is a Frege case along the L^* -route'. But there is a less stipulative-sounding formulation in the vicinity. Instead of explicitly ruling out the Frege-cases, we may impose that all members of the ordered set belonging to a single thinker be linked for that thinker. That is to say:

(Indirect Linking modulo FC) Two thoughts t_a and t_b stand in the indirect linking relation iff there is a tuple $\langle t_a, \dots, t_n, t_b \rangle$ such that:

- (i) each member stands in L to its successor;
- (ii)* all thoughts of a single thinker in $\langle t_a, \dots, t_n, t_b \rangle$ are linked for that thinker.

(ii)* essentially requires that any thoughts of a single thinker that are interpersonally linked should also be linked from the perspective of their thinker. Since **(IC)** defines sameness of thought in terms of linking, **(IC)** together with clause (ii)* validates a version of the *Transparency Constraint* for thought: for any two thoughts they deploy, a thinker should be in a position to know that the thoughts are the same, if they are the same.⁸ If we define Onofri's individuation criterion in terms of this definition of the indirect linking relation, the criterion is rendered compatible with **(FC)**.

As an upshot, the relation underwriting communicational success (that is, L) by itself does *not* ensure that the condition (ii)* will hold, on this proposal. Nor does the transitive closure of L (a different relation). Consequently, this raises the worry that shareability so construed is only an artificial construct which performs no genuine explanatory role (I won't discuss this worry here).

⁸See e.g. Boghossian (1994) for a discussion of this notion.

There are other ways to ensure that interpersonal linking is compatible with intrapersonal linking, so as to individuate shareable thoughts by the linking relation in a way that respects **(FC)**. An alternative is to add additional stringency to the *direct* linking relation itself (thus strengthening the condition required for successful communication, if one assumes that shareability explains communication). On this approach, the relation underwriting shareability *would*, in my opinion, have explanatory work to do, but at the cost of restricting the extension of what is shared and communicated⁹.

5 Conclusion

I have argued that Onofri's relational individuation criterion for thoughts does not satisfy Frege's constraint, because it identifies thoughts that are different for their thinkers. I proposed a way to fix the problem. But the problem is deeper: either the relational criterion must be of suitable granularity in the first place, or else it has to be explained how we can 'share' thoughts in a less-than-strict sense, if what we do is in fact less than strict same-thinking and same-saying.¹⁰

⁹For an explication of a notion of shareable content along these lines, at the cost of a much more stringent *same-content* relation, see Cumming (2013).

¹⁰I would like to thank Andreas Hallbäck, Michael Murez, and François Récanati for their helpful comments. This work was supported by the FrontCog funding under grant number ANR-17-EURE-0017.

References

- Boghossian, P. A. (1994). The transparency of mental content. *Philosophical Perspectives*, 8:33–50.
- Cappelen, H. and Hawthorne, J. (2009). *Relativism and Monadic Truth*, page 60. Oxford University Press UK.
- Cumming, S. (2013). Creatures of darkness. *Analytic Philosophy*, 54(4):379–400.
- Kripke, S. A. (1979). A puzzle about belief. In Margalit, A., editor, *Meaning and Use*, pages 239–83. Reidel.
- Kripke, S. A. (1980). *Naming and Necessity*. Harvard University Press.
- Onofri, A. (2018). The publicity of thought. *Philosophical Quarterly*, 68(272).
- Recanati, F. (2012). *Mental Files*. Oxford University Press.