



Itsetuntemus, ekspressivismi ja Wittgenstein

TERO VAAJA

Tämä artikkeli esittelee ekspressivististä itsetuntemuksen (*self-knowledge*) teoriaa erityyppisten itsetuntemuksen teorioiden kontekstissa.¹ Artikkelin viimekätinen tavoite on arvioida erityisesti ekspressivistisen teorian suhdetta Wittgensteinin mielenfilosofiaan, joka usein mainitaan teorian lähteenä, ja tarkastella teorian kehitystarpeita ja -mahdollisuuksia muiden itsetuntemuksen teorioiden valossa.

Osassa 1 esitetään eräitä käsitteellisiä selvennyksiä. Osa 2 käsittelee itsetuntemuksen (omia mentaalaisia tiloja koskevan tiedon) oletettuja filosofisesti kiinnostavia erityispiirteitä, joiden ymmärtäminen on itsetuntemuksen teorioiden arvioimisen kannalta olennaista. Osa 3 esittää itsetuntemuksen teorioiden karkean tyypittelyn sen mukaan, pitävätkö ne itseilmaisuja em-

¹ Artikkelin pohjautuu osittain Suomen Filosofisen Yhdistyksen kuu-kausikokouksessa ja Åbo Akademin tutkijaseminaarissa vuonna 2016 pitämiini esitelmiin. Kiitän molempien tilaisuuksien osanottajia lämpimästi keskusteluista ja kahta anonymia arvioitsijaa lukuisista tarkkanäköisistä ehdotuksista ja kommentteista.

piirisesti muodostettuina väittäminä, rationaaliseen toimijuuteen olennaisesti liittyvinä sitoutumina vai ekspressiivisinä lausumina. Ekspressivistisen vaihtoehdon yhteydessä arvioidaan sen suhdetta Wittgensteinin ajatteluun. Osassa 4 tarkastellaan Dorit Bar-Onin muotoilemaa neo-ekspressivististä itsetuntemuksen teoriaa ja esitetään, että ansioistaan huolimatta se ei ota tyydyttävästi huomioon rationaalisen toimijan kykyä konstituoida aktiivisesti ja reflektiivisesti ajatuksiaan ja asenteitaan. Lopuksi osassa 5 esitetään ajatuksia neo-ekspressivismin jatkokehittelyä varten ja kehoitetaan tarkastelemaan itseilmaisuja sosiallisissa kontekstissa Wittgensteinin kielipeliajatuksen mukaisesti.

1. Käsitteistöä

Harva aihe kuulostaa lähtökohtaisesti filosofisemmalta kuin itsetuntemus. Sellaisen itsetuntemuksen sijasta, jonka saavuttaminen on kiistatta vaikeaa – ymmärryksen omista kyvyistä, heikkouksista, vahvuuksista ja ajattelun vinoumista – analyttisen filosofian pääasiallisen argumentaation kohteena on kuitenkin ollut sen tyyppinen itseä koskeva tieto, jonka saavuttaminen vaikuttaa olevan helppoa ja yksinkertaista: tämänhetkisiä ajatuksia ja mielentiloja koskeva tieto. Tässä kirjoituksessa tuon argumentaation aihepiiristä käytetään termiä *itsetuntemus* englanninkielisen termin *self-knowledge* käännöksenä. On totta, että ”itsetuntemus” tuo terminä mieleen ensisijaisesti omien taipumusten, vahvuuksien ja heikkouksien tuntemisen, jota Cassam (2014) kutsuu ”substantiaaliseksi” itsetuntemukseksi erotettuna siitä ”triviaalista” (mutta tieto-opillisesti kiinnostavasta) omien ajatusten tietämisen kyvystä, johon analyttisen filosofian teoreettinen kiinnostus yleensä kohdistuu. Termin valinta on tehty ilmaisutavan yksinkertaistamisen nimissä.

Itsetuntemusta on luontevaa tarkastella ajatellen lausumia, joita tyyppillisesti kutsutaan vaikeasti käännettävällä termillä *avowals*. Termi on tuttu Ryleltä (1949/2009) ja Wittgensteinin teosten englanninkielisistä käännöksistä. Eräs tapa kääntää termi

olisi puhua *ensimmäisen persoonan preesens-muotoisista psykologisista väittämistä*. Pituuden ja kömpelyyden lisäksi on kuitenkin kaksi muutakin syytä olla käyttämättä tätä käännettä. Ensimmäinen syy, joka käy jatkossa seikkaperäisemmin selväksi, liittyy itsetuntemuksen teorioihin: niiden yhteydessä asettuu kyseenalaiseksi se, ovatko nämä lausumat luonteeltaan väittämiä vai toisen tyyppisiä puheakteja. Toinen syy on, että *avowal*-termillä halutaan tyypillisesti erottaa sellaiset lausumat, jotka edustavat puhujalle hänen varauksettomia ja välittömiä, rationaalisia, omiksi tunnustettuja mielentilojensa ilmaisuja. Kaikki ensimmäisen persoonan preesens-muotoiset psykologiset lausumat eivät ole tällaisia. Akraattinen tupakoitsija saattaa sanoa itsetutkiskelun hetkellä esimerkiksi: ”Haluaisin lopettaa tupakoinnin, mutta aina kun yritän, päädyn aloittamaan uudelleen. Tiedän, että tupakointi on pahaksi minulle ja lopettaminen olisi järkevää, mutta minulla on pakottava halu tupakoida.” Tarkastelun selkeyttämiseksi pilkotaan virke osiin seuraavalla tavalla:

- (1) Haluaisin lopettaa tupakoinnin
- (2) mutta aina kun yritän, päädyn aloittamaan uudelleen.
Tiedän, että tupakointi on pahaksi minulle ja lopettaminen olisi järkevää, mutta
- (3) minulla on pakottava halu tupakoida.

Avowal-termiä käytetään usein tavalla, jonka alaan lause (3) ei kuuluisi. Ajatuksena on, että akraattisen tupakoitsijan arvoskelma pakottavasta halustaan tupakoida on erityyppinen häntä koskeva psykologinen lausuma kuin hänen ilmaisunsa halustaan lopettaa, siis lause (1). Tupakoitsija muistelee ehkä heikko- tahtoista käytöstään pitkältä ajalta ja päätyy ilmaisemaan lauseella (3), että hänellä on pysyvä taipumus tupakanhimoon. Toisaalta lause (3) voi ilmaista myös tupakoitsijan tämänhetkisen, episodiluonteisen pakottavan halun tupakoida, jonka hän kuitenkin kokee ikään kuin ulkopuolisena pakottavana voimana: se on ristiriidassa hänen rationaalisesti muodostetun halunsa kanssa (1). Kummassakin lausuman (3) käyttötavassa tu-

pakoitsija tarkastelee itseään ikään kuin ketä tahansa ulkopuolista ihmistä, joka on objektiivisen psykologisen arvioinnin kohteena. Hän lukee halun tupakoida itselleen psykologisena faktana, mutta ei hyväksy tai omaksu sitä samassa mielessä kuin halunsa lopettaa. Tämä tärkeäksi koettu ero ilmenee terminologiassa niin, että *avowal*-nimitys halutaan tyypillisesti säästää lausumille, jotka puhuja esittää välittöminä mielentilansa ilmauksina, ottamatta itseensä teoreettista ja objektisoivaa asennetta.

Tässä kirjoituksessa *avowal*-termin suomenkieliseksi käännökseksi on valittu *itseilmaus*. Tämä ei tietenkään ole täydellinen käännös. *Avowal* on tyypillisesti sanallinen ilmaus, kun taas kaikki itseilmaisun muodot eivät tietenkään ole sanallisia. Myös sanallisia itseilmaisuja on monenlaisia: raivokas sadattelu ilmaisee kiukkua sanomatta verbaalisesti mitään sadattelijan mielentilasta. Jatkossa itseilmaisulla viitataan sellaisiin lausumiin, joissa puhuja ilmaisee mielentilansa eksplisiittisesti mentaalista käsitteistöä käyttäen – ”Tahdon lopettaa tupakoinnin”, ”Minulla on kipuja”, ”Uskon, että *p*” ja niin edelleen. Nämä moniselitteisyydet eivät toivottavasti aiheuta sekaannusta.

2. Explananda: itseä koskevan tiedon erikoispiirteet

Itseilmaisulla katsotaan tyypillisesti olevan joukko erityisiä tieto-opillisia ominaisuuksia, jotka näitä lausumia koskevan teorian tulisi kyetä selittämään. Olennaisia kandidaatteja ovat ainakin seuraavat kolme.

2.1 Ainutlaatuinen kanava

Itseilmaisujen mielentiloistamme välittämä tieto on peräisin lähteestä, joka on laadullisesti erilainen kuin se, jonka avulla muodostamme tietomme toisten ihmisten mielentiloista. Tiedon saamiseen toisten mielentiloista liittyy aina tavalla tai toi-

sella aistihavainto – toisen ihmisen verbaalisen ja kehollisen toiminnan havainnointi ja siitä enemmän tai vähemmän välittömästi tehty tulkinta. Ensimmäisen persoonan tapauksessa asia on kokonaan toisin. Minun ei tarvitse odottaa ja havainnoida omaa toimintaani tietääkseni, että uskon Tapparän liigamestaruuteen, että haluan pizzaa illalliseksi tai että vasenta säärtäni särkee. Tieto näistä asioista on peräisin erityisestä ensimmäisen persoonan kanavasta. Eräs itsetuntemuksen teorioita luokitteleva piirre on se, missä määrin havaintotiedon kaltaiseksi ne katsovat tämän ensimmäisen persoonan kanavan kautta tulevan tiedon, sikäli kuin kanava on olemassa. Rylen *The Concept of Mind* (1949/2009) on perusesimerkki kannasta, joka kieltää kanavan kokonaan ja väittää, että tietomme omista mielentiloistamme on pohjimmiltaan saman laatuista kuin tietomme toisten vastaavista mielentiloista. Rylen mukaan ainoa tieto-opillinen etu, joka itsetuntemukseen liittyy toisia koskevaan tietoon verrattuna, on itseämme koskevan havaintotiedon verrattoman suuri määrä. Ainutlaatuisen kanavan tunnustaminen explanandumiksi merkitsee, että riippumatta muista meriiteistään Rylen teoria oletetaan vääräksi ainakin tässä kohtaa.

2.2 Perustattomuus

Itseilmaisut vaikuttavat tieto-opillisesti perustattomilta. Kuten todettua, tietoni uskomuksestani, halustani tai jalkakivustani ei edellytä, että olisin johtanut tämän tiedon jostakin päättelyperustasta, esimerkiksi oman käyttäytymiseni havainnoinnista. Lisäksi näyttää siltä, että en voi (eikä minun tarvitse) esittää omaa mieltäni koskevien väitteiden tueksi mitään, mitä voisi kutsua episteemiseksi perusteeksi. Aistieni todistus on tietyissä mielessä peruste tiedolleni siitä, että keittiössäni on neljä tuolia. Voinee kuitenkin sanoa, että aistieni todistus on pätevä episteeminen peruste vain sen nojalla, että keittiön tuolien määrä on objektiivisesti tarkistettavissa oleva asiointila, jonka avulla aistieni yleistä luotettavuutta voidaan kontrolloida. Sen sijaan kanava, joka tuottaa tietomme omista mielistämme, vaikuttaa subjektiiviselta ja sen luotettavuus mahdottomalta tarkastella

intersubjektiivisesti. Silti olemme tyypillisesti valmiita hyväksymään toistemme tyypilliset itseilmaisut kyseenalaistamatta. Ilmaistessani liigamestaruutta koskevan uskomukseni, illallistoiveeni tai jalkasärkyeni, toisen henkilön vastaaminen kysymyksellä ”Mistä tiedät, että uskot/haluat/tunnet kipua?” vaikuttaisi kummalliselta ja ehkä tarkoituksettomalta. Tällainen kyseenalaistaminen tuntuu sotivan jonkinlaista normia vastaan.

Minkä tahansa (tai ainakin lähes minkä tahansa) muuntyyppisten tietoväitteiden kohdalla kyvyttömyys vastata kysymykseen ”Mistä tiedät?” muuten kuin sanomalla ”Minä vain tiedän” osoittaa vähintään jonkin asteista episteemisen velvollisuuden täyttämättä jäämistä. Poikkeuksia ovat mahdollisesti sellaiset välttämättömiä totuuksia ilmaisevat väitteet, joiden ymmärtämisestä seuraa välittömästi niiden todeksi tietäminen – mikäli tällaisia välttämättömiä totuuksia on olemassa. Loogiset ja matemaattiset totuudet voisivat ehkä olla esimerkkejä tapauksista, joita koskeva tieto ei edellytä muita perusteita kuin sen osoittamisen, että tietäjä on ymmärtänyt relevantit väitteet. Ehkä samanlaisen ominaisuuden voisi väittää olevan ilmeisillä moraalisisilla totuuksilla, kuten ”Ihmisten satuttaminen huvia vuoksi on väärin”. Itseilmaisut koskevat kuitenkin ilmiselvästi kontingenteja asiain-tiloja – ilmaisijansa ajatuksia, asenteita ja tunteita, joilla ei vaikuta olevan mitään tekemistä välttämättömien totuuksien kanssa. Itseilmaukset ovat syystä tai toisesta poikkeus säännöstä, jonka mukaan tieto (ainakin) kontingenteista asiain-tiloista edellyttää perusteita. Tällöin nousee kysymys, missä mielessä itseilmaisut lopultakaan ilmaisevat tietoa.

2.3 Ensimmäisen persoonan auktoriteetti

Itseilmaisut vaikuttavat olevan erityisen hyvin suojassa virheiltiltä. Se, että jokainen meistä on paras vastaamaan omia mielentilojaan koskeviin kysymyksiin, on osa tyypillisesti julkilauseumatonta arkiymmärrystämme. Mielen tiedostamattomien elementtien, itsepetoksen ja vastaavien seikkojen olemassaolo tietenkin osoittaa, ettei oma mieli ole kuitenkaan subjektille täysin läpinäkyvä. Meillä on toisinaan harhaisia tai vinoutuneita

käsityksiä esimerkiksi haluistamme, uskomuksistamme ja intentioistamme. Ensimmäisen persoonan auktoriteettia voi kuitenkin puolustaa maltillisemmassa mielessä kuin täydellisenä infallibilismina. Vaikuttaa kiistämättömältä, että kunkin henkilön itseilmaisulla on oikeutettu erityisasema tuon henkilön mielentilaa koskevien väitteiden joukossa. Paradoksaalista kyllä, vaikka emme yleensä edellytä kenenkään pystyvän esittämään episteemisiä perusteita itseilmaisulleen, emme pidä niitä lähtökohtaisesti kyseenalaisina, vaan lähtökohtaisesti tosina. Systemaattinen toisen ihmisen itseilmaisujen virheellisinä pitäminen olisi asiaankuulumatonta. Erityisesti tietyn tyyppinen virhe, havaintovirheeseen vertautuva ”puhdas erehdys”, näyttää olevan poissuljettu itseilmaisujen kohdalla (Burge 1996). Moitteettoman rationaalinen ihminen voi olla elimellisistä syistä systemaattisesti kykenemätön erottelamaan tiettyjä empiirisen todellisuuden kohteita, ja täten epäluotettava tiedon tuottaja niiden suhteen – esimerkiksi värisokea. Shoemakerin (1994) tunnetun argumentin mukaan tässä omaa mieltä koskeva tieto eroaa olennaisesti havaintotiedosta. Shoemakerin mukaan ei ole ajateltavissa, että rationaalinen toimija olisi vastaavalla tavalla ”itsesokea”, siis systemaattisesti kyvytön toimimaan luotettavana omien mielentilojensa ilmaisijana.

Kaikkiin edellä mainittuihin kolmeen piirteeseen on kuitenkin suhtauduttava varauksella. Jokaisen kohdalla on mahdollista väittää piirteen olevan vain filosofinen fiktio, tai että ainoastaan heikko versio siitä pitää paikkansa, tai että piirre (vahvana tai heikkona versiona) pätee ainoastaan rajalliseen tapausten joukkoon. Seuraavassa on lyhyesti joitakin huomioita. Kehityspsykologian valossa on selvää, että pikkulapsen ”mielen teoria”, eli kyky ymmärtää toisten subjektiivisia uskomuksia ja muita mentaalisia tiloja, kehittyy asteittain. Eräät havainnot viittaavat siihen, että kyky tehdä onnistuneita arvostelmia *omista* uskomuksista on myös vastaavalla tavalla kehittyvä ja teoreettinen (McGeer 1996; Gopnik & Astington 1988; Wimmer & Hartl 1991). Peter Carruthers (2011) on ansiokkaasti argumentoinut empiirisen todistusaineiston osoittavan, että myös aikuisten ar-

vostelmat omista uskomuksistaan ovat teoriapitoisia ja virhealttiita. Carruthersin mukaan paras selitys empiiriselle datalle on, että samat sosiaalisen kognition mekanismit (ns. *mind-reading*-järjestelmä), jotka mahdollistavat toisten ihmisten ymmärtämisen halujen ja uskomusten termein, toimivat myös silloin, kun attribuoiimme uskomuksia itsellemme. Tulkitsemme siis omia uskomuksiamme toimintamme perusteella enemmän tai vähemmän samoin kuin tulkitsemme toisten uskomuksia heidän toimintansa perusteella. Carruthersin mukaan olemme myös uskomustemme itsetulkinnassa alttiina aivan vastaavanlaisille tulkintavirheille kuin toisten ihmisten kohdalla. Jos Carruthers on oikeassa, on olemassa ainakin tapausten ala, jonka suhteen Ryle osui vähintään hyvin lähelle totuutta. Itsetuntemuksemme ei ole ainakaan kokonaan peräisin uniikista ensimmäisen persoonan kanavasta.

Edellä mainitut näkökohdat asettavat siis kyseenalaiseksi ajatuksen itsetuntemuksen ainutlaatuisesta episteemisestä kanavasta, mutta samalla myös sen perustattomuudesta ja auktoriteetista. Mikäli tietomme omista uskomuksistaamme perustuu samanlaiseen (paljolti tiedostamattomaan) kognitiiviseen prosessiin kuin toisten ihmisten käyttäytymisen tulkinta, se ei selvästikään ole episteemisesti perustatonta. Se, että emme tavallisesti *pysty vastaamaan* kysymykseen ”Mistä tiedät?”, ei tarkoita, etteikö kysymykseen olisi vastausta. Ja jos olemme systemaattisesti alttiita tietynlaisille virheille omien ajatustemme tulkinnan suhteen – vastaavanlaisille virheille, joita teemme toisten ihmisten tapauksessa – itseilmaisujen ensimmäisen persoonan auktoriteetti asettuu kyseenalaiseksi.

Tarkastelemaamme kolmiosaiseen explanandumiin tulee siis suhtautua varauksin ja ehdollisesti. Silti se puolustaa paikkaansa jossain määrin. Ensinnäkin lähtökohtainen taipumus ajatella itseilmaisujen olevan erikoislaatuisia enemmän tai vähemmän kolmen mainitun periaatteen mukaisesti, oli tämä sitten totta tai ei, on itsessään selitystä kaipaava ilmiö. Tämä taipumus on pätenyt ainakin mielenfilosofian tradition suhteen. Ei liene myöskään epäuskottavaa sanoa Wrightin (2015) tapaan, että filosofian lisäksi myös arkiajattelun mukaan:

[M]eistä jokainen on omiin mielentiloihimme erityisessä, toisilta suljetussa suhteessa [ja] merkittävään osaan mentaalista tiloistamme ja ominaisuuksistamme meillä itselämme on suora pääsy, toisilla vain epäsuora [...]. (Wright 2015, 49.)

Toiseksi, vaikka myönnettäisiin, että tietomme (esimerkiksi) omista uskomuksistamme ja haluistamme on välittyntä ja tietoisesta tai tiedostamattomasta itsetulkinnan tulosta, näyttää olevan mentaalisten tilojen luokka, jonka suhteen vastaava argumentti on vähemmän uskottava. Tähän luokkaan kuuluvat tuntemukset ja aistivaikutelmat kuten kivut ja näkö-, kuulo- ja proprioseptiset elämykset. Näitä koskevia arvostelmia voidaan karkeasti kutsua fenomenaaliksi arvostelmiksi. Fenomenaalisten arvostelmien auktoriteettia on vaikea kiistää. Millä tavalla sitten tarkkaan ottaen muodostammekin itseämme koskevia fenomenaalisia arvostelmia, ainakin sen tapa vaikuttaa tyystin erilaiselta kuin toisten ihmisten tapauksessa. Vaikka ajateltaisiin, että fenomenaaliset arvostelmat ovat erikoistapaus ja vaativat erilaista tieto-opillista käsittelyä kuin propositionaaliset asenteemme, niiden olemassaolo ainakin muistuttaa siitä, että itsetuntemuksen teorioiden on pakko olla sensitiivisiä eri mentaalisten tilojen tyyppien suhteen.

3. Mitä itseilmaisut ovat?

Filosofisia itsetuntemuksen teorioita voidaan karkeasti tyypitellä sen mukaan, millainen tehtävä itseilmaisulla niiden mukaan on inhimillisessä ajattelussa ja kommunikaatiossa. Tähän liittyvät läheisesti seuraavat kysymykset. Ovatko itseilmaisut yleisesti ottaen samantyyppisiä totuuskelpoisia väitteitä kuin kontingentteja maailman asiaintiloja koskevat lausumat yleensä? Sikäli kuin itseilmaisut ovat aitoja (ja auktoriteettia nauttavia) tietoväitteitä mielentiloistamme, millaista niiden episteeminen oikeutus on luonteeltaan? Yksinkertaisuuden nimissä kysymys voidaan asettaa seuraavasti: minkätyyppisiä puheakteja itseilmaisut (pääasiassa tai ensisijaisesti) ovat?

3.1 Empiirisiä väittämiä?

Ensimmäinen vastausvaihtoehto on yksinkertaisin ja ilmeisin: itseilmaisut ovat empiirisiä väittämiä. Ne ovat deskriptiivisiä lausumia, joiden kohteena on oma mieleemme, ja tässä suhteessa samalla tasolla toisten mieliä koskevien väittämien kanssa. Naturalistiseen kuvaan mielestä voi kuulua oletus kognitiivisesta itsetarkkailutoiminnosta, joka monitoroi subpersonaalisia mielentiloja ja tekee ne tietoisiksi. Sikäli kuin itseilmaisujen oletetaan olevan tällaisen itsemonitoroinnin tuloksena syntyviä enemmän tai vähemmän tarkkoja raportteja puhujan mielentiloista, ne ymmärretään empiirisiksi väitteiksi. Näin ollen itseilmaisut eivät olisi laadultaan mitenkään olennaisesti erilaisia aistihavaintoon perustuvien empiiristen väitteiden kanssa. Niiden syntymekanismi olisi toki erilainen kuin esimerkiksi näkö- tai kuulohavaintoraporttien, mutta syvempää laadullista eroa tapausten välillä ei olisi. Olisi tietenkin naiivia ajatella, että saamme tietää omat ajatuksemme ”kuuntelemalla sisäistä puhettamme” muuten kuin korkeintaan erikoistapauksissa, mutta yleisesti ottaen itsetuntemuksemme lähteenä voitaisiin ajatella olevan kyky (mahdollisesti se, jota kutsumme introspektioksi), joka on luonteeltaan havaintokykyyn verrattava.

Itseilmaisujen voidaan ajatella olevan luonteeltaan empiirisiä väitteitä myös pitämättä niitä introspektion tuotteina (Gertler 2016). Voidaan kuitenkin esimerkinomaisesti tarkastella kantaa, jonka mukaan itseilmaisut ovat yksinkertaisesti eräs empiiristen tosiasiaväitteiden laji, jonka episteeminen oikeutus perustuu introspektioon vastaavasti kuin visuaaliset arvostelmat perustuvat näköaistin todistukseen. Tällöin kolmi-osainen explanandum näyttäytyisi seuraavasti. Itseilmaisut perustuvat ainutlaatuisen tiedolliseen kanavaan: Tämä kanava on pohjimmiltaan Locken (1689/2000) ”sisäisen aistin” kaltainen ja sitä kutsutaan introspektioksi. Itseilmaisut eivät ole todella perustattomia, vaan niiden episteeminen perusta on niille ominainen havaintokyky samalla tavalla kuin näkökyky on visuaalisten arvostelmien episteeminen perusta. Taipumus ajatella toisin johtuu vain siitä, että introspektiomme tapahtuu yleisimmin tiedostamattomasti. Meillä on itseilmaisuidemme

suhteen ensimmäisen persoonan auktoriteetti siksi, että introspektiivinen pääsy mielentiloihin on tietenkin vain kulloistenkin mielentilojen subjektilla. Kukin on yksinään ainutlaatuisella näköalapaikalla omaan mieleensä, joten kullakin on mielensä suhteen automaattisesti episteeminen etulyöntiasema. Tämä kuitenkin tekee itseilmaisusta vain – jossain määrin – epäämättömiä (*incorrigible*), ei erehtymättömiä (*infallible*). Muut ihmiset ovat aina lähtökohtaisesti heikossa asemassa korjaamaan itseilmaisujani, mutta mitään syvällistä takuuta niiden totuudesta ei sinänsä ole.

Ajatukseen itseilmaisusta empiirisinä väittäminä liittyy kuitenkin ongelmia. Ensinnäkin, empiirisiin väittämiin liittyvät olennaisesti *väittämän kohteen ja havainnointiprosessin erillisyy*s (Shoemaker 1994) sekä *havaintovirheen mahdollisuus*. Fenomenaalisten arvostelmien kohdalla ajatus on ongelmallinen. Mikäli esimerkiksi kipuelämyksen ajatellaan olevan olemassa subpersonaalisena mentaalisenä tilana riippumatta siitä, tunnistaako itsemonitorointitoimintomme sen, tällöin omista kipuelämyksistä tulisi olla mahdollista erehtyä samalla tavalla kuin mistä tahansa muustakin havaintotiedon kohteesta. Tämä ei kuitenkaan ole selvää. Ajatellaan henkilöä, jonka keho ja keskushermosto ovat tilassa, johon tavallisesti liittyy kipuelämys, mutta kyseinen muuten kognitiivisesti normaali henkilö on kykenemätön havaitsemaan kipuelämystään ja raportoimaan sitä (koska itsemonitorointitoiminto syystä tai toisesta pettää). Tällöin kysymys siitä, tuntee ko hän kipua vai ei, olisi itse kivun käsitteen nojalla moniselitteinen. Kipu on keskeisessä merkityksessään tietoinen elämys, jonka osittain konstituoivat sen kognitiivinen ”käsilläolo” ja raportoitavuus. Vaikka tämä ei olisi täysin kiistaton käsitteellinen totuus kivun suhteen, se lienee ainakin tarpeeksi uskottava kipukäsitteen prima facie -käytön kuvauksena, että se tulee ottaa vakavasti.

Toisentyyppinen ongelma liittyy uskomusten ilmaisuihin. Uskomuksia koskevien itseilmaisujen ymmärtäminen luonteeltaan empiirisiksi nostaa kysymyksiä näiden suhteesta rationaalisuuteen ja toimijuuteen. Oman rationaalisen ajattelun suhteen

näyttää olevan olemassa *toimijaperspektiivin* ja *havaittajaperspektiivin* välinen tärkeä ero, joka empiristisen kannan tulisi selittää. Eroa havainnollistaa seuraava alun perin Tyler Burgelta peräisin oleva esimerkki (Gertler 2016):

Kuvitellaan kaksi tuomaria, jotka uskovat vastaajan syyllisyyteen. *Harkitseva tuomari* tietää uskovansa vastaajan syyllisyyteen. Punnittuaan todistusaineistoa huolellisesti hän tulee kuitenkin siihen tulokseen, että todisteet tämän syyllisyyden puolesta ovat hyvin heikkoja. Päätelmä aiheuttaa välittömästi sen, että harkitseva tuomari hylkää alkuperäisen uskomuksensa vastaajan syyllisyydestä. Hän tietää, ettei nyt usko vastaajan syyllisyyteen, koska todistusaineisto pakottaa rationaalisesti tämän mielenmuutoksen hänelle. Toinen tuomari, *pakkomielteinen tuomari*, tietää myös uskovansa vastaajan syyllisyyteen. Hän tutustuu samaan todistusaineistoon ja niin ikään toteaa todisteiden olevan hyvin heikkoja. Pakkomielteinen tuomari kuitenkin havainnoi itseään introspektiivisesti ja toteaa uskovansa yhä vastaajan syyllisyyteen. Uskomus säilyy, vaikka pakkomielteinen tuomari tietää sen olevan todisteiden valossa epärationaalinen. Myös pakkomielteinen tuomari ehkä päätyy lopulta hylkäämään uskomuksensa, mutta joutuu vapauttamaan itsensä uskomuksestaan epäsuorin keinoin – esimerkiksi kuvittelemalla toistuvasti vastaajan jonkin muualle kuin rikospaikalle tai keskittymällä intensiivisesti muiden epäiltyjen syyllisyyden puolesta puhuviin todisteisiin.

Pakkomielteinen tuomari tietää kylläkin uskomuksensa, mutta hänellä on siitä vain ”empiiristä”, tietyllä tavalla ulko-kohtaista tietoa. Harkitseva tuomari tietää uskomuksensa tavalla, joka ei ole luonteeltaan pelkästään empiirinen. Tämän kaltaisten esimerkkien on tarkoitus muistuttaa siitä, että asianmukainen rationaalinen suhde omiin uskomuksiimme vaatii muutakin kuin pelkästään sitä, että attribuimme niitä itsellemme (vaikka kuinka tarkkanäköisesti). Henkilö, joka on aidosti uskomustensa subjekti ja niistä rationaalisesti vastuussa, tietää uskomuksensa tyypillisemmin niiden *muodostajan* kuin havainnoijan ominaisuudessa.

3.2 Sitoutumia?

Kuten Shoemaker (1994) huomautti, kyky tietää luotettavasti omat uskomuksensa on rationaalisen toimijuuden ennakkoehto. Ellei tiedä, mitä uskoo, ei myöskään pysty huolehtimaan siitä, että omat uskomukset ovat uskollisia evidenssille rationaalisuuden vaatimusten mukaisesti (Burge 1996). Kykymme tietää uskomuksemme luotettavasti – josta seuraa ensimmäisen persoonan auktoriteettimme niiden suhteen – takaa se, että emme normaalioloissa vain huomaa uskovamme johonkin, vaan tietoisesti muodostamme uskomuksia aiemman tietomme ja evidenssin pohjalta reflektoiden. Uskomuksia koskevilla itseilmaisuuilla voi näin olla puheakteina sitoutumien (*commitments*), ei empiiristen väitteiden luonne.

Eräät itsetuntemuksen teoriat korostavat uskomusten ja noita uskomuksia koskevien toisen asteen asenteiden ei-kontingenttia yhteyttä. Näille teorioille on annettu nimi *konstitutioismi*. Ajatuksena on, että tiettyjä propositionaalisia asenteita (ainakin uskomuksia) koskevat ensimmäisen persoonan ylemmän tason uskomukset ja arvostelmat ovat (tyypillisesti) konstitutiivisia noille asenteille itselleen. Colivan (2016, 163–166) esitystä mukaillen, konstitutivisti pitää seuraavantyyppistä teesiä tiettyjen olosuhteiden vallitessa totena:

(KONS) A uskoo, että $P \leftrightarrow A$ uskoo uskovansa, että P

Teesi sanoo, että kun A:lla on uskomus P, tästä seuraa, että hänellä on myös toisen asteen uskomus. Lisäksi, kun A muodostaa uskomuksen tai esittää arvostelman, että hän uskoo P:n, tästä seuraa, että hän todella uskoo P:n.

Ilman täsmennyksiä KONS vaikuttaisi ilman muuta väärältä. Ihmisillä ei ole reflektiivistä suhdetta kaikkiin uskomuksiinsa: voimme uskoa monia asioita ilman, että koskaan muodostamme niistä toisen asteen uskomuksia. Luemme myös sujuvasti uskomuksia vastasyntyneille ja ei-ihmiseläimille, joilla ei oletettavasti ole reflektiokykyä. Toisaalta toisen asteen uskomuksen läsnäolo ei ole aina tae ensimmäisen asteen uskomuksen olemassaolosta. On mahdollista lukea itselleen uskomuksia,

joita ei todellisuudessa omaa. Saatan rehellisesti uskoa uskovan Tapparan liigamestaruuteen tulevalla kaudella, mutta silti lyödä vetoa JYPin mestaruuden puolesta. Kyseessä on ehkä toiveajattelu tai itsepetos, jolloin toisen asteen uskomukseni ei yksiselitteisesti ratkaise kysymystä, tuleeko minun sanoa uskoneen Tapparan liigamestaruuteen.

Tarvittavaksi täsmennykseksi voidaan Colivan (2009, 2016) ehdotuksen mukaisesti esittää seuraavaa. Millaisten olosuhteiden vallitessa KONS pitää paikkansa? Teesi on hyväksyttävä silloin, kun se pätee nimenomaan *sitoutuman* luonteisiin uskomuksiin, jotka ovat elimellinen osa inhimillistä rationaalisuutta. Colivan määritelmän mukaan sitoutumat ovat subjektin aktiivisesti muodostamia asenteita, jotka ovat tiedostetusti sensitiivisiä järkeilyn normatiivisille periaatteille. Näiden seikkojen vuoksi ne ovat asenteita, joista subjekti on rationaalisesti vastuussa (2016, 189).² Ei-ihmiseläimet ja vastasyntyneet eivät ole rationaalisia toimijoita eivätkä muodosta sitoutumaluonteisia uskomuksia, joten näiden uskomuksiin KONS ei päde. Sitoutumaluonteinen uskomus voi olla myös rinnakkain olemassa sen kanssa ristiriitaisen, dispositioluonteisen uskomuksen kanssa. Minulla on sitoutumaluonteinen uskomus Tapparan liigamestaruudesta sikäli kuin olen tietoisesti punninnut evidenssiä, todennut uskomuksen perustelluksi ja attribuoinut sen itselleni välittömästi tämän nojalla. Sitoutumaluonteista uskomustani ei kumoa se, että minulla voi samaan aikaan olla ei-rationaalinen uskomus JYPin mestaruudesta. Epärationaalisuuteni tässä asiassa ei murra sitoutumaluonteisen uskomukseni ja sitä koskevan toisen asteen arvostelman välttämätöntä yhteyttä.

² Colivan määritelmä kokonaisuudessaan kuuluu: Sitoutumat ovat propositionaalisia asenteita, jotka (i) seuraavat subjektin toiminnasta, nimittäin arvostelman muodostamisen mentaalisesta teosta, (ii) ovat teoreettisen ja käytännöllisen päättelyn periaatteiden normatiivisesti rajoittamia, (iii) ovat noiden periaatteiden rajoittamia myös subjektin omasta näkökulmasta ja (iv) mielentiloja, joista subjekti on vastuussa (2016, 189).

Ajatellaan nyt edeltävän kappaleen esimerkkiä harkitsevasta ja pakkomielteisestä tuomarista. Nähdään, että pakkomielteinen tuomari ei voisi säilyttää statustaan täydessä mitassa rationaalisen toimijana, ellei hän olisi kykenevä muodostamaan pakkomielteisen uskomuksensa vastapainoksi sitoutumaluonteista uskomusta vastaajan syyttömyydestä, joka ei ole empiirisesti tai introspektiivisesti löydetty vaan reflektiivisesti muodostettu. Tuomari motivoituu karistamaan pakkomielteisen uskomuksensa siksi, että hän kykenee tulemaan refleктоimalla tulokseen, että hänen ei *pitäisi* uskoa vastaajan syyllisyyteen – mikä implikoi suoraan sen, että hän *ei usko* vastaajan syyllisyyteen sitoutumamielessä.

Niissä itseilmaisun tapauksissa, joissa ilmaistaan sitoutumaluonteinen uskomus (esimerkiksi ”En usko (enää), että vastaaja on syyllinen”), kyseessä ei ole siis pelkästään empiirinen lausuma puhujan mentaalista tilasta. Lausuma ilmaisee rationaalisen harkintaprosessin lopputuleman, joka tekee konstitutiivisesti todeksi sen, että puhujalla on – sitoutumamielessä – ensimmäisen asteen uskomus, joka ilmaistaisiin sanoilla ”Vastaaja ei ole syyllinen”. Sikäli kuin lausuma on rehellinen, se on tosi väite puhujansa mielentilasta, mutta sen totuutta ei takaa puhujan introspektion luotettavuus, vaan lausuman asema osana hänen tiedostavaa toimintaansa järkeilijänä.

Rationaalisuuden ja itsetuntemuksen välinen keskeinen suhde ei koske pelkästään uskomuksia, vaan myös tunteita, asenteita ja muita intentionaalisia tiloja, joiden suhteen meillä on eettis-rationaalisia velvollisuuksia. Jotta pystyisin arvioimaan, ovatko esimerkiksi reaktioni epäoikeudenmukaisuuksiin oikeanlaisia ja oikein mitoitettuja, minun on pystyttävä sanomaan, missä määrin närkeä havaitessani epäoikeudenmukaisuutta, yli- tai alireagoinko, ja niin edelleen. Ellei varttuneilla ihmisillä olisi yleisesti ottaen kykyä tietää omat tunnetilansa luotettavasti, emme voisi pitää heitä täydessä mitassa kyvykkäinä moraalisena reflektion. Tällaisiin pohdintoihin perustuu itsetuntemuksen toimija- tai *agentiaalinen* näkemys, jonka mukaan itsetuntemuksen ja rationaalisen toimijuuden

kyvyt ovat yhteydessä toisiinsa. Tiedämme omat mielentilamme erityisellä tavalla, koska olemme itse niiden luojia ja ylläpitäjiä ja muokkaamme niitä sensitiivisesti perusteiden ja evidenssin suhteen.

Kaikki edellä sanottu on ehkä oikein, mutta sitä voidaan syyttää myös puheenaiheen vaihtamiseksi. Agentiaalinen teoria näyttää soveltuvan vain sellaiseen tapausten luokkaan, joka ei ole kovin keskeinen omaa mieltä koskevan tiedon episteemisten ongelmien suhteen. On totta, että ”saamme tietää” uskomuksemme tai asenteemme muotoillessamme ne tietoisesti ja sitoutuessamme niihin, mutta tämä on eri asia kuin jo valmiiksi olemassa olevan uskomuksen tai asenteen selville ottaminen. On mahdollista väittää, että itsetuntemuksen teorioiden kannalta kiinnostavia ovat ensisijaisesti ne tapaukset, joissa omaa mielentilaa koskevan harkinnan kohteena oleva tila on selkeästi erillinen harkintaprosessista itsestään. Tällöin olisi sanottava, että agentiaalinen teoria saattaa olla sinänsä pätevä ja artikuloida tärkeitä seikkoja rationaalisen toimijuuden ja itsereflektion suhteesta, mutta se ei varsinaisesti vastaa itsetuntemuksen kannalta olennaisiin episteemisiin kysymyksiin.

Ajatellaan seuraavaa havainnollistusta. Ajanhetkellä t_1 minulla ei ole mitään tietoisista asennoitumista kysymykseen, mitä haluan tehdä tänä iltana töistä päästyäni. Välittömästi tämän jälkeen, ajanhetkellä t_2 , joku esittää minulle kysymyksen, mitä haluan tehdä tänä iltana töistä päästyäni. Harkitsen asiaa ja muodostan joukon tietoisia uskomuksia. En ole pitkään aikaan jutellut ystäväni Miken kanssa; Mike todennäköisesti ilahtuisi, jos soittaisin hänelle; minusta olisi miellyttävää jutella Miken kanssa, ja niin edelleen. Tulen siis tietoiseksi halu-uskomus-yhdistelmästä, joka motivoi minut soittamaan Mikelle. Tämän seurauksena tulen siihen tulokseen, että haluan soittaa Mikelle töistä päästyäni. Nyt voin raportoida kysyjälle tämän juuri muodostuneen halun, ja tällöin samalla sitoudun toimimaan töistä päästyäni päätökseni mukaisesti, ellen halua toimintani olevan epärationaalista suhteessa itseäni ja Mikeä koskeviin tietoihin haluihini ja uskomuksiini. Ei ole vaikeaa nähdä, miksi raporttini halustani soittaa Mikelle on erityisen luotettavasti

tosin. Kyseisen halun muodostaminen oli kaiken aikaa minusta itsestäni kiinni, ilmaisemalla haluni julkisesti luon samalla itseleni lisämotivaation toimia sen mukaisesti, ja kun sitten toimin johdonmukaisesti ja todella soitan Mikelle töiden jälkeen, tämä vahvistaa raporttini olleen luotettava lausunto mielenliikkeistäni. Mutta kysymys siitä, mitä halusin aiemmin tietyllä ajanhetkellä – esimerkiksi ajanhetkellä t_1 välittömästi ennen harkintaprosessini alkua – näyttää olevan eri kysymys kuin se, mihin harkintani antaa vastauksen. Itse asiassa voisin tulkita keskustelukumppanini kysymyksen, mitä haluan tehdä töiden jälkeen, kahdella tavalla: joko kehotuksena harkita asiaa, tai kehotuksena spontaanisti raportoida, mitä sillä hetkellä haluan. Minulta vaadittavat kognitiiviset tehtävät olisivat tapauksissa erilaiset. Harkintani jälkeen en vaikuta olevan lainkaan paremmin varustautunut vastaamaan kysymykseen, mitä halusin tehdä töiden jälkeen ajanhetkellä t_1 (tai jos en halunnut erityisesti mitään, mistä tiedän *sen?*), tai miten tiesin ajanhetkellä t_2 niin ilmeisen varmasti esimerkiksi sen, että minusta olisi miellyttävää puhua Miken kanssa. Jälkimmäisten kaltaisiin episteemisiin kysymyksiin agentiaalinen teoria ei millään ilmeisellä tavalla vastaa.

Lisäksi on sanottava, että vaikka agentiaalinen teoria olisikin (edellä sanotusta huolimatta) valaiseva uskomusten, tunteiden ja asenteiden suhteen, sen ala rajoittuu joka tapauksessa vain ”perusteiden suhteen sensitiivisiin” mielentiloihin, eli mielentiloihin, joiden kohdalla on mielekästä puhua niiden muodostamisesta ja niihin sitoutumisesta. Se ei tarjoa uskottavaa vastausta kysymykseen, millaista on luonteeltaan esimerkiksi aistikokemuksia ja elämyksiä koskeva tieto, sikäli kuin nämä eivät ole missään olennaisessa mielessä subjektin tietoisesti muokkaamia tai muodostamia. Emme tavallisesti luo itsellemme elämyksiä ja aistikokemuksia, vaan ne tapahtuvat meille.

3.3 Ekspressiivisiä lausumia?

Kaikki itseilmaisut eivät ole rationaalisuuden kannalta sensitiivisiä tavalla, josta agentiaalinen näkemys puhuu. On olemassa

ainakin tapausten luokka, jossa tiedämme oman mentaalisen tilamme ilman reflektiivistä harkintaa tai itsetarkastelua, mutta kuitenkin niin, että erehtymisen mahdollisuutta on vaikea kuvitella. Aistielämysten ja monien tahtotilojen ilmaukset toimivat esimerkkeinä. Eräs reaktio tähän on omaksua kanta, joka usein yhdistetään Wittgensteinin myöhäiskauden mielenfilosofiaan. Tämän kannan mukaan itseilmaisujen erikoispiirteet selittyvät sillä, että ne eivät ole raportteja tai kuvauksia esittäjensä mielentiloista, vaan ekspressiivisiä lausumia. Itseilmaisut ovat vain näennäisesti väitelauseiden kaltaisia. Todellisuudessa niiden virka ei ole raportoida tai väittää mitään, vaan tuoda julki puhujansa tiloja ja tarpeita samaan tapaan kuin primitiiviset, eikielelliset ilmaisut tekevät. Tästä seuraa, että itseilmaisuille ei tulisi lukea totuusarvoja. Niiden välittömyys ja erehtymättömyys selittyvät deflationistisesti sillä, että myöntäminen, kieltäminen, perusteiden kysyminen ja epäileminen ovat niiden suhteen asiaankuulumattomia reaktioita. Ne olisivat asiaankuulumattomia reaktioita itkemiseen surun ilmauksena, ja yksinkertaisen ekspressivismin kannattajalle sama pätee itseilmauksiin ylipäänsä. Wittgenstein kirjoittaa tunnetussa katkelmassa:

Miten sanat viittaavat aistimukseen? [...] Kysymys on sama kuin: Miten ihminen oppii aistimusten nimien merkityksen? Esimerkiksi sanan "kipu" merkityksen. Tämä on yksi mahdollisuus: Sanat yhdistetään aistimuksen alkuperäiseen, luonnolliseen ilmaukseen ja sanoja käytetään tämän ilmauksen paikalla. Lapsi on loukannut itsensä ja huutaa. Aikuiset puhuvat sitten hänelle ja välittävät hänelle huudahduksia ja myöhemmin lauseita. He opettavat lapselle uuden kipukäyttäytymisen. (Wittgenstein 1953/1981, §244.)

"Minuun sattuu" on lausuma, joka alkaa kielen oppimisen myötä toimittaa samaa virkaa kuin kivusta valittaminen. Pikkulapsi oppii kurottelemisen sijasta sanomaan: "Tahtoo pallon." Näin päädytään yksinkertaiseen *ekspressivistiseen* itseilmaisujen

teoriaan, jonka mukaan itseilmaisut eivät kuvaa (sisäistä tai ulkoista) maailmaa, vaan pelkästään ilmaisevat puhujansa tunteita ja tarpeita. Teoria siis käsittelee itseilmaisuja samalla tavalla kuin metaeettinen ekspressivismi moraalilausumia. Se myös törmää samaan vaikeuteen, niin sanottuun Frege-Geach-ongelmaan: Jos itseilmaisut ovat pelkästään ekspressiivisiä, miten selitetään se, että ne näyttävät kuitenkin toimivan rationaalisen päättelyn osina kuten väitelauseet? Seuraava päätelmä näyttää ilman muuta pätevältä:

[1] Olen Helsingissä

[2] Minuun sattuu

Siis [3] Helsingissä on ainakin yksi henkilö, johon sattuu.

Jos yksinkertainen ekspressivistinen teoria pitää paikkansa, toisella premissillä ei ole totuusarvotettavaa sisältöä, eikä päättelyn tulisi vaikuttaa näin ollen pätevältä. Tämä on kuitenkin järjenvastaista. Jollei halua hylätä oletusta tämän kaltaisen päättelyn pätevyydestä, on myönnettävä, että itseilmaisuuilla on ekspressiivisen käyttönsä rinnalla vähintään vaihtoehtoinen deskriptiivinen käyttö – minkä muun muassa Wittgenstein (1953/1981, ss. osa II, ix) suorasanaisesti myöntää. Deskriptiivisesti käytettynä ”Minuun sattuu” voi olla tosi tai epätosi kuvaus sisäisestä tilastani ja toimia siten myös päättelyn premisinä. Tällöin nousee kuitenkin välittömästi kysymys, millaista väitteinä käytettyjen itseilmaisujen episteeminen oikeuttaminen on ja mihin niiden ensimmäisen persoonan auktoriteetti perustuu.

Yksinkertainen ekspressivismi ostaa ratkaisunsa itseilmaisujen erikoispiirteiden selittämiseen liian kalliilla hinnalla. Se jättää kokonaan huomiotta sen, että käytämme itseilmaisuja kiistämättä (myös) deskriptiivis-rationaalisesti. Tuloksena on absurdi kuva itseilmaisujen luonteesta ja roolista kommunikaatiossa. Teoria johtaa muun muassa siihen, että jos henkilö A.B. sanoo ”N.N.:ään sattuu”, ja tämän jälkeen N.N. itse sanoo ”Kyllä, minuun sattuu”, tällöin N.N. ei ole vahvistanut A.B.:n väittämää. Itse asiassa mikään, mitä N.N. voi sanoa, ei ole periaatteessakaan laskettavissa saman, A.B.:n kanssa jaetun

väittämän esittämiseksi. Näin ollen näyttää välttämättömältä antaa itseilmaisuja koskevalle ekspressivismille jokin sellainen muoto, joka pystyy selittämään, miksi itseilmaisulla on mahdollista käydä – niiden ekspressiivisyydestä huolimatta – tosiasiaväitteiden esittämisen ja mentaalisten asiointilojen kuvaamisen kaltaista diskurssia.³

Crispin Wright (1998) on esittänyt, että itseilmaisujen erikoispiirteitä ei tulisi pyrkiä selittämään millään niiden luonnetta ja semantiikkaa koskevalla teorialla, vaan niiden käyttöä ohjaavien kielellisten konventioiden kautta. Wrightin mukaan tämä oli myös Wittgensteinin todellinen kanta. Wrightin mukaan Wittgensteinin ei pidä katsoa sitoutuneen ekspressivistiseen teoriaan, jota Wright pitää toivottomana. Hänen mukaansa Wittgensteinin keskeinen väite oli, että filosofinen teoria, joka pyrkii selittämään itseilmaisujen erikoispiirteet, hahmottaa selittäjän ja selitettävän väärin päin. Itseilmaisulla ei ole erikoispiirteitä siksi, että ne perustuvat subjektin ainutlaatuiseseen, välittömään suhteeseen omaan mentaaliseen maailmaansa. Sen sijaan se, että itseilmaisuja käytetään ja niihin suhtaudutaan kielessä tietyllä tavalla, osaltaan määrittää sitä, mistä puhumme ”mentaalisisä ilmiöinä”. Se, että itseilmaisuilta ei edellytetä epistemisiä perusteita eikä niiden oikeellisuutta tavallisesti haasteta, on osa kielellistä käytäntöä – osa sitä, mistä Wittgenstein puhui ”mentaalisten termien kielioppina”. Wrightin mukaan itseilmaisujen kielelliseen käytäntöön kuuluu, että niiden hyväksyminen tosiksi on ”default-vaihtoehto”

³ Tarvitaan siis kehitelmä, jonka suhde yksinkertaiseen itseilmaisuja koskevaan ekspressivismiin olisi jotakuinkin sama kuin metaeettisen kvasi-realismiin (Blackburn 2006, 2007) suhde yksinkertaiseen metaeettiseen ekspressivismiin. Jäljempänä tarkasteltu neo-ekspressivistinen teoria on lupaava ehdotus ainakin joissakin suhteissa tällaiseksi kehitelmäksi. Neo-ekspressivistisen teorian ja metaeettisen kvasi-realismien välinen analogia ansaitsisi oman lähemmän käsittelynsä. Tämän artikkelin puitteissa tyydyttävä tarkastelu ei kuitenkaan ole mahdollista.

– aina lähtökohtaisesti oikein, ellei jokin tiedonintressi erityisesti edellytä henkilön itseilmaisujen asettamista kyseenalaiseksi ja teoreettisen näkökulman ottamista. Näin tapahtuu esimerkiksi psykiatrisissa terapiaistunnoissa.

Wrightin mielestä Wittgensteinin viesti itseilmaisujen filosofian suhteen oli tämän ajattelulle tyypillisesti selitysvastainen. Wittgenstein epäili ilmiön selittämiseen pyrkivien teorioiden käyttökelpoisuutta, ei sitä, etteikö itseilmaisuja voisi käyttää väitelauseina ja lukea niille totuusarvoja. Wright toteaa, että ekspressivismin ei edes tarvitsisi sitoutua kategoriseen jakoon ekspressiivisten lausumien ja väitelauseiden välille sekä itseilmaisujen totuusarvottomuuteen. Hienovaraisempi ekspressivistinen teoria voisi sanoa, että itseilmaisujen ensisijainen käytötapa on ekspressiivinen, mutta tämän sivutuotteena ne voivat saada myös väitelausemaisia käyttöyhteyksiä ja totuusarvoja. Sama pätee performatiivien kohdalla: ”Julistan teidät mieheksi ja vaimoksi” on ensisijaiselta käyttötavaltaan julistus, ei puhujan senhetkistä toimintaa koskeva väite. Tämä ei sulje pois sitä, että se *samalla* kuvaa kyseisen toiminnan, ja voi olla tosi tai epätosi riippuen siitä, täyttyvätkö ensisijaisen puheaktin suorittamisen ehdot.

Wright hylkää kuitenkin myös hienovaraisemman ekspressivistisen vaihtoehdon liian ongelmallisena. Hänen ensisijainen perusteensa on, että teoria ei myöntäisi subjekteilla olevan mitään episteemistä erityissuhdetta omiin mielentiloihinsa, mikä olisi Wrightin mukaan absurdia. Hän on ehkä tässä kohtaa liian hätäinen. Hänen oma näkemyksensä (jonka hän lukee myös Wittgensteinille) on yhtä lailla ongelmallinen. Kuten edellä todettiin, itseilmaisujen tiedollinen erikoislaatuisuus liittyy läheisesti rationaaliseen toimijuuteen. Se, onko subjekti täydessä mitassa rationaalinen toimija, on osaltaan riippuvainen siitä, osaako hän kompetentisti tehdä selkoa sisäisestä toiminnastaan ja elämyksistään. Tämä ei voi olla pelkästään sosiaalisesta hyväksynnästä riippuva asia. Henkilön, jonka itseilmaisuihin on oikein ottaa perusteita kysymätön ja niiden auktoriteettia kunnioittava asenne, on täytynyt onnistua muussakin kuin siinä, että saa ympäröivän kieliyhteisön hyväksymään lausumansa.

Statusta itsetuntemukseen kykenevänä subjektina ei voi vain saada, se täytyy ansaita – vaikka tämä ansaitseminen olisi luonteeltaan hyvinkin minimaalista. Kuten Wright myöntää, hänen wittgensteinilaisen ratkaisunsa hyväksyminen edellyttää lisäksi, että hyväksymme wittgensteinilaisen käsityksen filosofisen selittämisen rajoista, ja tämä metafilosofinen keskustelu ei ratkea millään yksinkertaisella tavalla.

4. Neo-ekspressivismi

Ekspressivististä itseilmaisujen teoriaa on sittemmin kehitetty. Dorit Bar-On (2004) ja David Finkelstein (2003) ovat esittäneet kannat, joita voidaan kutsua Bar-Onin termin mukaisesti *neo-ekspressivistisiksi*. Bar-Onin versio itseilmaisuja koskevasta ekspressivismistä myöntää, että itseilmaisut voivat olla väitteitä ihmisten arvostelmista ja ilmaisukyvystä riippumattomasti olemassa olevista mentaalisisistä tosiseikoista ja että itseilmaisuille voidaan lukea totuusarvoja. Näin ollen Bar-Onin neo-ekspressivismi ei ole yksinkertaisen metaeettisen ekspressivismin kaltainen kanta, joka laskisi alaansa kuuluvat lausumat vain näennäisesti tosiseikkoja koskeviksi väitteiksi. Bar-On kuitenkin säilyttää ekspressivismin ydinsisältönä pitämänsä kannan, jonka mukaan itseilmaisut eivät ole luonteeltaan mentaalisen tilan kuvauksia. Ne eivät ole sanoiksi käännettyjä raportteja subjektin yksityisesti havainnoimasta mentaalisestä tilasta, vaan noiden tilojen kognitiivisesti välittömiä julkilausumia. Bar-Onin tavoite on siis esittää omaperäinen, eri suuntaisia intuitioita kunnioittava itsetuntemuksen teoria yhdistämällä kaksi vastakkaisilta vaikuttavaa ajatusta: itseilmaisut ovat sekä totuusarvolettavia väitteitä että ekspressiivisiä ja ei-kuvaavia. Teorian ytimessä on ajatus, että sanalliset itseilmaisut ovat ekspressiivisiä *tekoja*, jotka kuitenkin semanttisessa mielessä ilmaisevat propositioneja. Bar-On pyrkii selittämään itseilmaisujen ekspressiivisen luonteen avulla, kuinka ensimmäisen persoonan auktoriteetti voidaan ymmärtää niin, että itseilmaisut ovat tavallisesti kirjaimellisesti *tosia*. Kanta on siis tavallaan tinkimättömämpi

kuin yksinkertaisen ekspressivismin ehdotus, joka suo itseilmaisuille epäämättömyyden, mutta vie niiltä totuusarvot.

Bar-Onin mukaan itseilmaisut kuten ”Minuun sattuu” tai ”Tahdon ulos” *ilmaisevat toiminnallisesti* tiettyjä mielentiloja samassa merkityksessä kuin vastaava luonnollinen toiminta – kivusta valittaminen tai oven hakkaaminen. Toiminnallinen ilmaiseminen on luonteeltaan kognitiivisesti välitöntä. Bar-Onin keskeinen väite on, että samaan tapaan kuin ”minä”-pronominin käyttö ei edellytä käyttäjältään kykyä tunnistaa itseään missään episteemisesti sisällöllisessä mielessä, myöskään itseilmaisut eivät edellytä käyttäjältään niitä vastaavien sisäisten tilojen tunnistamista. Ajatellaan tilannetta, jossa olen ilmapiiriltään epämiellyttävässä kokouksessa, ja lopulta sanon: ”Nyt riittää – tahdon ulos.” Sikäli kuin minulla on perusteita sanomalleni, perusteet ovat samoja kuin mitä esittäisin kellon vilkuilemiselle, turhautuneiden katseiden luomiselle ovea kohti ja muille sanattomille toiminnallisille ilmauksille. Perusteeni voisivat olla ne, että kokousta ei käydä rakentavassa ja kunnioittavassa hengessä, että näissä olosuhteissa pidetty kokous on hyödytön ja että tilanne vaatii passiivisen protestin esittämistä. Perusteisiin ei kuuluisi se, että olisin introspektiivisesti havainnut itsessäni mentaalisen tilan ja tunnistanut sen haluksi lähteä ulos. Voin olla väärässä perusteissani (ehkä olen tulkinnut kokoustajien tarkoitusperät väärin tai ylireagoanut), mutta minun ei voida sanoa erehtyvän oman mielentilani tunnistamisessa, koska mielentilan tunnistamisella ei ole tilanteessa lainkaan kognitiivista roolia. En erehdy siinä, koska en tee sitä alun alkaenkaan.

Bar-On välttää yksinkertaisen ekspressivismin epäuskottavat seuraukset tekemällä eron toiminnallisen ilmaisemisen ja *semanttisen ilmaisemisen* välillä. Lausumani ”Tahdon ulos” on ekspressiivinen akti, joka ilmaisee tahtotilani toiminnallisesti, mutta samalla se ilmaisee semanttisesti erään propositionaalisen sisällön. Tuon semanttisesti ilmaistun sisällön nojalla lausumani on myös totuusarvotettavissa ja käytettävissä loogisesti väitelauseen tapaan. (Bar-On 2009, 67–70.)

Bar-Onin kehitelty ekspressivismin versio vaikuttaa lupavalta, sikäli kuin sen tarkoitus on tehdä selkoa artikkelin alussa

eritellystä kolmiosaisesta explanandumista. Neo-ekspressivismin mukaan sanalliset itseilmaisumme ovat toiminnallisen ilmaisukykykymme tuotteita. Ne eivät perustu oman käyttäytymisemme havainnointiin (eivätkä introspektioon). Tämä käy yksiin sen ajatuksen kanssa, että itseilmaisut tuottaa ainutlaatuisen, olennaisesti subjektiivinen kanava. Itseilmaisut ovat myös episteemisesti perustattomia siinä mielessä, että itseilmaisulleni mahdollisesti antamani perusteet koskevat ainoastaan toiminnallisen ilmaisun aktini oikeellisuutta. Kokousesimerkissä minulta voidaan kysyä, miksi tahdon ulos, ja voin esittää asialle perusteet. Minulta voidaan myös kysyä: ”Tahdotko todella ulos?” tai: ”Oletko varma, että tahdot ulos?” Nämä kysymykset eivät kuitenkaan kohdistu siihen, olenko tunnistanut mielentilani oikein – tai jos kohdistuvat, niitä on oikeutettua pitää illegiitimienä kysymyksinä. Ne ovat vain painokkaampia tapoja kyseenalaistaa toiminnallisen ilmaisuni perusteet. Itseilmaisuille kuuluu ensimmäisen persoonan auktoriteetti, koska niitä ei voida mielekkäästi haastaa episteemisesti samaan tapaan kuin tavallisia tosiasiaväitteitä, vaan ainoastaan toiminnallisen itseilmaisun oikeellisuuden osalta kulloisessakin kontekstissa.

Bar-Onin mukaan itseilmaisujen auktoriteetti ei ole ehdoton. Toisinaan toiminnalliset ilmaisumme tuottavat semanttisina ilmaisuinaan epätosia propositioita. Bar-On pitää tätä teorian sa vahvuutena, koska hänen mukaansa neo-ekspressivistinen teoria selittää itseilmaisujen lähtökohtaisen auktoriteetin, mutta sallii niiden olevan tietyssä – arkijärjen mukaisessa – mitassa alttiita virheille. Hän käyttää seuraavanlaista esimerkkiä kuvaamaan tilannetta, jossa yksinkertainen rehellinen kivun ilmaus voi olla epätosi. Potilas istuu hammaslääkärin tuolissa ja odottaa tuntevansa pian kipua. Pora lähestyy hammasta, ja ennen sen osumista hampaaseen potilas vavahtaa ja huutaa: ”Satuu!” Ilmaisun on välitön ja ekspressiivinen, ja potilaan voidaan sanoa uskovan sen todeksi. Se on kuitenkin kiistatta epätosi. Toiminnallinen ilmaus voi epäonnistua tavalla, jonka tuloksena syntyy epätosi semanttinen ilmaisu. Potilas erehtyy, mutta Bar-On korostaa, että tämä erehtyminen on luonteeltaan erilaatuista kuin havaintoarvostelmassa erehtyminen (Bar-On 2009, 57).

Omasta kivustaan voi tässä mielessä erehtyä, koska kivut ovat ontologisesti erillisiä niitä koskevista arvostelmista ja asenteista. Rehellinen kivunilmaisuni kertoo tyypillisesti, että minulla on kipuja, mutta ei implikoi sitä – *“saying so does not make it so”*. Kaikki mentaaliset tilat eivät kuitenkaan ole vastaavalla tavalla ontologisesti itsenäisiä. Sellaisten sitoutumaluontoisten halujen (ja uskomusten), jollaisia käsitellään yllä osassa 3, olemassaolo riippuu niitä koskevista asenteistani. Sitoutumaluonteinen haluni soittaa tänä iltana Mikelle on olemassa vain sikäli kuin se on itselleni lukema tietoinen halu. Itseilmaukseni *“Haluan soittaa tänä iltana Mikelle”* (tai valmiuteni esittää se tarvittaessa toisille tai ajatuksissa itselleni) konstituoii osaltaan sitoutumaluonteista haluani. Halun olemassaolo *on* valmiuttani lukea kyseinen halu itselleni ja toimia sen mukaisesti.

Bar-On tekee selväksi, että haluaa kunnioittaa *Riippumattomuudeksi* kutsumaansa periaatetta: hän haluaa teoriansa pätevän mentaalisiin tiloihin, jotka eivät ole ontologisesti arvostelmista tai asenteista riippuvaisia (Bar-On 2009, 56–57). Hän ei siis hyväksy itseilmaisujen auktoriteetin selittämistä tavalla, joka tekisi mentaalisisista tiloista yleisemmin ottaen sitoutumien kaltaisia. Neo-ekspressivismissä itseilmaisun ja ilmaistavan välinen suhde ei ole konstitutiivinen.

Millainen tuo suhde sitten on? Ennen kaikkea, mikä sulkee pois sen, etteivät subjektin rehelliset itseilmaukset voisi jatkuvasti täysin epäonnistua heijastelemaan hänen mielentilojaan, joita niiden oli määrä ilmaista? Kuten näimme, agentiaalisen teorian vastalause empirismille oli, että asianmukainen rationaalinen suhde omiin ajatuksiimme ja asenteisiimme edellyttää, että emme saa tietää niitä pelkästään introspektion tai vastaavan itsehavainnointiprosessin kautta, vaan niiden aktiivisen muodostajan roolissa. Eräitä mielentilojamme koskevat toisen asteen asenteemme vaikuttavat olevan konstitutiivisia niiden mielentilojen itsensä suhteen. Jos neo-ekspressivistinen teoria ei tunnusta tällaista suhdetta, se on altis samalle agentialistiselle kritiikille kuin empiristinen teoria. Empiristisen teorian mukaan itsetuntemuksessa on kyse itsehavainnoinnista, jonka an-

siosta voimme raportoida tuohon havainnointiin nähden ontologisesti riippumattoman mentaalisen tilan; neo-ekspressivismin mukaan kyse on itseilmauksesta, joka ilmaisee tuosta ilmaisun aktista riippumattoman mentaalisen tilan. Ellei itseilmauksen ja ilmaistavan välillä ole mitään apriorista konstituu-tio- tai vastaavaa suhdetta, voitaisiin niin empiristisen kuin neo-ekspressivistisenkin teorian valossa kuvitella itesokea subjekti, jonka rehelliset itseilmaukset tuottavat jatkuvasti epä-tosia väittämiä. Subjektin toiminnallinen itseilmaisukyky voisi periaatteessa olla täysin epäluotettava hänen omia mielentilo-jaan koskevan tiedon tuottaja. Tämä vaikuttaa kuitenkin para-doksaaliselta. Se jättää huomiotta agentialistisen kritiikin, jonka mukaan rationaalisen ajattelijan toisen asteen asenteet ovat yh-teydessä hänen ensimmäisen asteen mielentiloihinsa tavalla, joka tekee täydellisestä itesokeudesta periaatteessakin mah-dottoman. Ei ole selvää, että neo-ekspressivismi pystyisi vastaa-maan agentialistiseen kritiikkiin lainkaan empirististä teoriaa paremmin. Tällöin uhkaa näyttää siltä, että myös neo-ekspres-sivistinen teoria on joko väärä tai ei sovellu laajaan ja tärkeään joukkoon tapauksia.

Mikäli itseilmaisuja ajatellaan ainoastaan ne synnyttävien mielentilojen tyypillisesti luotettavina indikaattoreina, vaikut-taa jossain määrin ristiriitaiselta kuvata niitä *ekspressiivisiksi*. Kun itseilmaisusta puhutaan esimerkiksi taiteellisen ilmaisun kohdalla, on tyypillistä ajatella, että ilmaisuprosessi on sekoitus luomista ja löytämistä, jossa ilmaistava ajatus tai tunne ei ole täysin erotettavissa ilmaisutavasta. Jos neo-ekspressivistinen teoria ei tunnusta vastavuoroista suhdetta itseilmaisun ja il-maistavan välillä, tämä saattaa johtaa siihen, että jotain olen-naista ensimmäisen persoonan auktoriteetin luonteesta jää se-littämättä.

5. Itseilmaisuu, autenttisuus ja sitoutuminen

Tarkastellaan tätä kysymystä palaamalla kokousesimerkkiin. Neo-ekspressivististä teoriaa voidaan kehittää käsitevälinein,

joiden aihiot ovat ainakin jossain määrin löydettävissä Wittgensteinilta.

Kuvitellaan, että jälkeinpäin minut hyvin tunteva henkilö kommentoi käytöstäni kokouksessa seuraavasti: ”En epäile, ettekö olisi ollut rehellinen sanoissasi, mutta tahdoitko todella ulos? Tiedämme molemmat, että et ole konflikteja vierastava persoona. Nautit väittelyistä etkä yleensä halua perääntyä kiistatilanteista. Näit varmasti, kuinka laitosjohtaja oli silmin nähden närkästynyt ihmisten käytöksestä. Tiedän, että haluat näyttää hyvältä laitosjohtajan silmissä. Olen varma, että todellisuudessa olisit halunnut jäädä – tahdoit lähteä vain myötäilläksesi pomon mielialoja.”

Ystäväni kyseenalaistaa tässä itseilmaisuni totuusarvon kyseenalaistamatta sanomani rehellisyyttä. Hän ei syytä minua valheesta, mutta syyttää itseilmaisua *epäautenttisuudesta*. Hän ehdottaa, että toimintani ja sanomani ei heijastellut todellisia taipumuksiani ja ajatuksiani, vaan ulkoapäin omaksumiani sosiaalisia normeja tai tiedostamattomia käyttäytymismalleja.

Voin reagoida tähän useilla eri tavoilla. Yksi tapa on suora kieltäminen: saatan kieltää perusteet, joille ystäväni pohjaa epäautenttisuusväitteensä. Toisaalta saatan hyväksyä hänen perusteensa, mutta väittää, että niistä huolimatta todella halusin lähteä kokouksesta, koska ilmaisemani tahtotila oli luonteeltaan sitoutuma. Väitän, että lähtöpäätökseni oli tietoisin harkinnan tulos: arvioin tilanteen vaatimuksia, ehkä tiedostaen sekä oman houkutuseni jäädä jatkamaan stimuloivaa väittelyä että haluni esiintyä edukseni laitosjohtajan silmissä, ja tulin siihen tulokseen, että kaikki asiat huomioon ottaen paras toimintatapa oli lähteä ulos ja sitoutua tähän päätökseen lausumalla se julkisesti. Harry Frankfurtin (1971) terminologian hengessä tilanteen voisi kuvata niin, että tiedostin ensimmäisen asteen haluni jatkaa väittelyä, mutta sitouduin sen kumoavaan toisen asteen haluun. Väitän tällöin siis, että sitoutumiseni toisen asteen haluun oli konstitutiivista sen suhteen, mikä tahtotilani tuolla hetkellä todella oli. Kolmanneksi, saatan hyväksyä ystäväni näkemyksen ja päätyä ajattelemaan, että tahdonilmaukseeni todella vai-

kuttivat tiedostamatta jääneet tekijät, joiden huomaamisen jälkeen en ole enää tyytyväinen lausumaani tahtoni ilmauksena. Saatan joko päätyä pitämään itseilmaisuaani epätotena tai omak-sua epävakaa asenteen sitä kohtaan, horjuen sitoutuman ja epäautenttinen ilmaus -käsitteiden välillä.⁴

Ei ole mitään ilmeistä syytä olettaa, että jokin vastausvaihtoehdoista vastaa tosiseikkoja paremmin kuin toiset. Valinta saattaa olla yksinkertaisesti alideterminoitu kaikkiin todisteisiin nähden, mukaan luettuna oman introspektioni ”todistus”. Eräs aspekti Wittgensteinin kritiikistä sisäisen kokemuksen reifiointia kohtaan (”sisäinen tapahtuma’ tarvitsee ulkoiset kriteerit” (Wittgenstein 1953/1981, §580)) on tässä relevantti. On paikallaan huomata, että kaikki se, mitä kutsumme kuvaamiseksi, ei ole välttämättä ”sanojen lukemista tosiseikoista” – sitä, että kuvaamme ”tosiseikat sanoiksi sääntöjen mukaan” (Wittgenstein 1953/1981, §292). Se, että jotakin lausumaa, esimerkiksi ”Tahdon ulos”, voidaan käsitellä mielentilani kuvauksena, ei sulje pois sitä, että se voi samalla olla aktiivisesti ja tietoisesti muodostamani sitoutuma, välitön spontaani tunteen ilmaus tai statukseltaan epämääräisesti näiden välissä. Mielentilani kuvaus voi monissa tapauksissa olla luonteeltaan senhetkisen parhaan itseymmärrykseni varauksellinen julkituominen, jonka lopulli-

⁴ Burgen (1996) käyttämien termien mukaisesti voitaisiin sanoa, että valinnassa on kyse siitä, päädyntkö pitämään näkökulmaani omaan mieleeni tilanteessa yhtenäisenä vai epäyhtenäisenä. Epäyhtenäisyys tarkoittaisi tässä sitä, että tunnustan itsessäni kirjaimellisesti kaksi toistensa kanssa ristiriitaista halua: halun jäädä kokoukseen ja halun poistua sieltä. Burge käyttää termiä *point of view*: Minulla on tällöin yksi näkökulma, josta katsottuna haluan jäädä kokoukseen, ja toinen näkökulma, josta katsottuna haluan lähteä. Jos käsittelen lausumaani sitoutumana, tämä tarkoittaa, että tunnustan itsessäni vain yhden, yhtenäisen näkökulman, ja haluan yksiselitteisesti lähteä. Tällöin olen alentanut vastakkaisen halun silkaksi (tavallaan eläimelliseksi) ensimmäisen asteen haluksi, joka ei edusta minulle ”näkökulmaa” itseäni.

nen hyväksyttävyyys – sekä omissani että toisten silmissä – riippuu viime kädessä siitä, kuinka ymmärrettävästi ja johdonmukaisesti se asettuu laajempaan kuvaan minusta persoonana.

Se, mihin itsetulkintaan päädyn, riippuu reflektiivisen järjestyksen prosessista, johon ottavat osaa myös toiset ihmiset. Sitä voidaan osuvasti kuvailla wittgensteinilaisella kielipelin metaforalla. Itseilmaisuni autenttisuuden haastaminen, ilmaisuni auktoriteetista kiinni pitäminen ja ilmaisuni tulkitseminen sitoutumaksi ovat tämän erityisen kielipelin ”siirtoja”. Kuten kielipelin ideaan kuuluu, peli voi olla luonteeltaan kilpailullinen (keskustelu voi olla muodoltaan *väittely* siitä, mitä todella halusin), mutta sen ei tarvitse olla. Ensimmäisen persoonan auktoriteetin voi tällöin ymmärtää olevan eräänlainen kielellisen ja sosiaalisen varttumisen myötä kehittyvä kyky. Se on kykyä osallistua toisten ihmisten kanssa jaettuun itsetulkinnan prosessiin ja praktista ymmärrystä siitä, koska ja millä perusteilla minun on oikeutettua väittää tiettyjen sanojeni tai toimieni konstituoivan sen, mitä todella ajattelen. Paras tapa kehittää itsetuntemuksen teoriaa ei ehkä olekaan pyrkiä saattamaan ilmiön kaikki muodot yhden homogeenisen teorian alle, vaan tunnustaa tiedonmuodostusprosessin olevan eri tapauksissa erilainen, ja tutkia erilaisten oman mielensä tuntemisen tapojen soveltamisaloja ja keskinäisiä riippuvuussuhteita.

Neo-ekspressivistisen teorian kehittäminen tähän tapaan tarjoaisi rikkaamman kuvauksen itseilmaisun ja ilmaistavan välisestä suhteesta. Se myös selittäisi sen, miksi täydellinen itsesokeus vaikuttaa järjenvastaiselta (koska subjekteilla katsotaan olevan tiettyyn mittaan asti rationaalinen auktoriteetti väittää, että jotkut heidän itseilmaisuihistaan ovat konstitutiivisia heidän ajatuksiensa ja asenteidensa suhteen), mutta myös pitäisi kiinni siitä validista ajatuksesta, että itseilmaisujen auktoriteetti ei ole ehdoton. Bar-On ei ehkä kuitenkaan hyväksyisi tätä kehittelyä, koska se edellyttäisi Riippumattomuusperiaatteen ainakin jonkin asteista hylkäämistä. Toisaalta Bar-On kirjoittaa myös, ettei hän sulje pois ensimmäisen persoonan auktoriteetin voivan olla jonkin tyyppisten ilmaisujen kohdalla peräisin myös konstitutiivisesta suhteesta (2009, 63n21). Näin ollen ei ole välttämättä

liian radikaali ehdotus, että sitoutuman luonteisilla itseilmaisuuilla voi olla paikka myös kattavammassa itsetuntemuksen teoriassa, joka säilyttää neo-ekspressivismin perusajatuksen.

Wrightin mukaan Wittgenstein ei halunnut selittää itseilmaisujen erityisyyttä millään perustavanlaatuisemmilla kieltä tai ajattelua koskevilla faktoilla. Näin ollen Wittgensteinia ei tulisi pitää myöskään ekspressivistinä – tai ylipäänsä minkään teorian kannattajana, joka katsoo itseilmaisuuilla olevan jokin olemuksellinen semanttinen luonne. Nähdäkseni Wittgenstein voisi kuitenkin hyväksyä edellä kuvatun kaltaisen neo-ekspressivismin kehittelyn. Edellä kuvattua teoriaa voidaan pitää luonteeltaan ”yleiskatsauksena” tai itseilmaisun käytäntöjen käsitteellistämisen tapana, jonka filosofinen elinvoima riippuu viime kädessä siitä, kuinka valaiseva se on tutkimamme ”kielipelin” käytännön ymmärtämisessä.

Jyväskylän yliopisto

Kirjallisuus

- Bar-On, D. (2004). *Speaking My Mind: Expression and Self-Knowledge*. Oxford: Oxford University Press.
- Bar-On, D. (2009). “First-Person Authority: Dualism, Constitutivism, and Neo-Expressivism”, *Erkenntnis* 71 (1), 53–71.
- Blackburn, S. (2006). “The Semantics of Non-factualism, Non-cognitivism, and Quasi-realism”, teoksessa M. Devitt & R. Hanley (toim.), *The Blackwell Guide to the Philosophy of Language*. Oxford: Blackwell, 244–251.
- Blackburn, S. (2007). “Antirealist Expressivism and Quasi-Realism”, teoksessa D. Copp (toim.), *The Oxford Handbook of Ethical Theory*. New York, Oxford: Oxford University Press, 146–161.
- Burge, T. (1996). “Our Entitlement to Self-Knowledge”, *Proceedings of the Aristotelian Society, New Series*, 96: 91–116.
- Carruthers, P. (2011). *The Opacity of Mind: an integrative theory of self-knowledge*. Oxford: Oxford University Press.
- Cassam, Q. (2014). *Self-Knowledge for Humans*. Oxford: Oxford University Press.
- Coliva, A. (2009). “Self-Knowledge and Commitments”, *Synthese*, 171 (3), 365–375.

- Coliva, A. (2016). *The Varieties of Self-Knowledge*. Palgrave Macmillan UK.
- Finkelstein, D. (2003). *Expression and the Inner*. Cambridge: Harvard University Press.
- Frankfurt, H. G. (1971). "Freedom of the Will and the Concept of a Person", *The Journal of Philosophy*, 68 (1), 5–20.
- Gertler, B. (2016). "Self-Knowledge and Rational Agency: A Defense of Empiricism", *Philosophy and Phenomenological Research*, julkaistu verkossa, DOI: 10.1111/phpr.12288
- Gopnik, A. & Astington, J. W. (1988). "Children's Understanding of Representational Change and Its Relation to the Understanding of False Belief and the Appearance-Reality Distinction", *Child Development* 59 (1), 26–37.
- Locke, J. (1689/2000). *An Essay Concerning Human Understanding*. Kitchener, Ont.: Batoche Books.
- McGeer, V. (1996). "Is 'Self-Knowledge' An Empirical Problem? Renegotiating the Space of Philosophical Explanation", *The Journal of Philosophy* 94 (10), 483–515.
- Ryle, G. (1949/2009). *The Concept of Mind*. Lontoo: Routledge.
- Shoemaker, S. (1994). "Self-Knowledge and 'Inner Sense': Lecture I: The Object Perception Model", *Philosophy and Phenomenological Research* 54 (2), 249–269.
- Wimmer, H. & Hartl, M. (1991). "Against the Cartesian View on Mind: Young Children's Difficulty with Own False Beliefs", *British Journal of Developmental Psychology*, 9, 125–138.
- Wittgenstein, L. (1953/1981). *Filosofisia tutkimuksia*. Juva: WSOY.
- Wright, C. (1998). "Self-Knowledge: The Wittgensteinian Legacy", *Royal Institute of Philosophy Supplement* 43, 101–122.
- Wright, C. (2015). "Self-Knowledge: The Reality of Privileged Access", teoksessa S. Goldberg (toim.), *Externalism, Self-Knowledge and Scepticism: New Essays*, Cambridge: Cambridge University Press, 49–74.

