Coevolution of human speech and trade

Richard D. Horan · Erwin H. Bulte · Jason F. Shogren

Abstract   We propose a paleoeconomic coevolutionary explanation for the origin of speech in modern humans. The coevolutionary process, in which trade facilitates speech and speech facilitates trade, gives rise to multiple stable trajectories. While a 'trade-speech' equilibrium is not an inevitable outcome for modern humans, we find it is a relatively likely scenario given our species evolved in Africa under climatic conditions supporting relatively high population densities. The origin of human speech is not independent of economic institutions—the economics of early human trade can provide additional insight to help explain the physiological emergence of human speech.

Keywords   Evolution · Language · Coordination game · Early humans · Paleoeconomics

JEL Classifications   N50 · O13 · Q20 · Q57

1 Introduction

Human speech allows for superior communication and is one of the primary reasons our species has been so successful. But why the morphology for human speech evolved remains

Richard D. Horan (✉)
Department of Agricultural Economics, Michigan State University, East Lansing, MI 48824-1039, USA

Erwin H. Bulte
Development Economics Group, Wageningen University, Wageningen, The Netherlands

Erwin H. Bulte
Department of Economics, Tilburg University, P.O. Box 8130, Wageningen, 6700 EW, The Netherlands

Jason F. Shogren
Department of Economics and Finance, University of Wyoming, Laramie, WY 82071-3985, USA

a mystery (Holden 1998).[1] Most agree the explanation must be partly behavioral, as the evolutionary benefits of human speech (i.e., improved communication) primarily stem from improved social interactions (Knight et al. 2000). But what behaviors would have made the morphological investments worthwhile, and how did this evolutionary process occur?[2] In this paper we develop a *paleoeconomic* theory of the coevolution of trade and speech, whereby the tradeoffs associated with each investment jointly influence the evolutionary and behavioral outcomes.[3]

Investments in human speech morphology were costly. These investments, which were in place between 150,000 and 200,000 B.P. (Lieberman 1998; Holden 1998), involved changes in the upper respiratory system—specifically the supralaryngeal vocal tract (SVT). One change was a descended larynx, which increased choking risks. Previously humans could breathe and drink or consume small amounts of food simultaneously, without choking (Lieberman 1984). Another change involved the loss of a "snout", which increased the curvature of the airway and reduced aerobic efficiency (Lieberman 1984). Lieberman (Liemberman 2007) indicates that only humans (*Homo sapiens*) have a human SVT; all other hominids had a "standard plan" SVT, resulting in vocal limitations relative to *Homo sapiens*.[4]

What behaviors could have led to this investment? Several theories exist, though detailed evidence of hominid behaviors from the period of morphological evolution is difficult to unearth (Holden 1998). A popular theory is *vocal grooming*. This theory is based on the observation that primates "groom" each other to help support relationships, which suggests speech could enable larger, more stable communities by allowing each individual to "vocally" groom more people (Dunbar 1996). But others counter that vocal grooming is "cheap" relative to physical grooming (or other forms of non-vocal signals) and may be unreliable (see Knight et al. 2000; Dessalles 2000). Moreover, this and other theories based on cooperation

---

[1] There is also a question of *when* speech emerged. Unequivocal evidence of spoken language does not exist until about 40,000 B.P., long after the morphology for human speech was in place (Holden 1998). Lieberman 1984, Liemberman 2007) argues this gap between morphology and cultural evidence is a chimera—the morphology had to evolve for speech because it was simply too costly otherwise. Others agree that speech is a biological adaptation (Pinker and Bloom 1990), arguing that our present day inability to find convincing historical evidence best explains this time lag (Holden 1998). Corballis (2002) suggests the morphology could have initially been used to produce songs useful in some sexual selection process, and that speech emerged later. But he acknowledges (as do many others) that proto-speech would have been used in conjunction with hand and facial gestures during this morphological evolution.

[2] Investments in human speech required cognitive and morphological adaptations. Many cognitive adaptations came first. Paleoneurologists indicate the Broca's and Wernicke's areas of the brain, which are now important in language production, processing and comprehension, evolved as far back as *Homo habilis* (2.5–2 million B.P.; Wilkins and Wakefield 1995). It is believed these areas evolved to support conceptual abilities associated with tool making and use, and gestural communication (Paillard 2000; Wilkins and Wakefield 1995). Though their initial purpose was not to support speech, these and other cognitive adaptations eventually would have supported primitive speech (Paillard 2000; Wilkins and Wakefield 1995), referred to as proto-speech, as early as 500,000–1.5 million B.P. (Knight et al. 2000). These adaptations also formed the neural preconditions for more complex speech (Paillard 2000; Wilkins and Wakefield 1995). We take earlier cognitive adaptations and primitive speech abilities as given and focus on morphological adaptations.

[3] Within the scope of the paleoeconomics field are fundamental scientific mysteries like the evolution of humans and human behaviour (Hansson and Stuart 1990; Galor and Moav 2002; Robson and Kaplan 2003; Rogers 1984), early human resource management and food procurement strategies (Smith 1975; Brander and Taylor 1998; **?**; Bulte et al. 2006; Marceau and Myers 2006; Weisdorf 2005); the importance of biogeography (Olsson and Hibbs 2005); and the emergence of trade, early markets and property rights (Ofek 2001; Lagerlöf 2005). The common process of these models is to highlight behavioral explanations by introducing formal economic modeling into research fields like anthropology and archaeology traditionally dominated by more descriptive approaches.

[4] The extent of these limits cannot be known entirely without more information about the shape and descent of the non-human species' tongues (Liemberman 2007).

do not address the co-evolution of the behavior and human speech morphology.[5] That is, they do not consider morphological changes and the associated physiological costs; rather they focus on the selection of speech or language as a behavior, implicitly assuming the morphology is already in place.

More recently, Ofek (2001) and Robson (2005) have described how increased biological complexity, such as human speech morphology, can arise due to a Red Queen game—a coevolutionary arms race between competitors seeking an evolutionary advantage. Ofek (2001) discusses how human speech might arise from a *verbal arms race* driven by a uniquely human activity—competition for trading partners in a market exchange system (i.e., exchange between non-family members). Here the best communicators gain the relative trading advantage. Robson's (2005) social intelligence theory features within-species competition as the key co-evolutionary factor in the verbal arms race (as does Ofek).

But could competition for trading partners have been the driving force? Not necessarily. Some evidence exists suggesting that trading networks among early modern humans existed by at least 130,000 B.P. (Holden 1998), and Ofek (2001) provides convincing arguments that it occurred much sooner—perhaps 1.5–2 million B.P.[6] This suggests hominids were already trading by the time the morphological adaptations for speech were underway. But the data is insufficient to suggest the widespread trading needed to generate significant competition for trading partners. More likely, trading was emerging during this time frame and the difficulty was not getting someone to trade with you over someone else, but rather to get anyone to trade with you at all. Emerging trading systems can be modeled as a cooperative activity in which trading only emerges through coordination (e.g., Williamson and Wright 1993).

When trade is modeled as a cooperative activity, multiple equilibria are possible due to strategic complementarities—more trade by one agent increases another agent's gains from trade. Hominids must coordinate their decisions to capture gains from trade that outweigh the associated transactions costs, such as those arising from communication barriers. Coordination succeeds under the right bioeconomic circumstances and sufficient expectations for this outcome (Diamond 1982; Williamson and Wright 1993). Coordination failure, however, arises when the trading equilibrium fails to emerge due to insufficient expectations (Cooper and John 1988).

When viewed in a coevolutionary framework, the results for our model in this paper suggest the gains from trade can make nature's investments in communication worthwhile, which further increases the gains from trade. But human speech and trade are not guaranteed, as multiple equilibria may arise due to potential coordination failures. Bioeconomic conditions,

---

[5] Some scholars have suggested that the believability of signals can be improved via costly rituals (Knight 1998; see also Noble 2000), social contracts (Deacon 1997; see also Gifford 2002) or rules of kin selection or spatial selection (Livingstone and Fyfe 2000). Another theory is that speech evolved as a knowledge transfer mechanism (Pinker and Bloom 1990), though free-riding, deception, and cheating could make reliance on second-hand information an unreliable strategy for evolutionary success (e.g.,Knight et al. 2000). A third cooperative theory is based on production—that speech emerged, though possibly after the morphology was in place, to free up hands otherwise used to communicate through gesture or to converse in the dark (e.g., Corballis 2002). The literature on the evolution of altruistic behaviors and social institutions is substantial; see Bowles et al. (2003) for a recent analysis.

[6] Ofek (2001) notes that members of the *Homo* lineage are the only primates to abandon the "feed-as-you-go" strategy and bring food back to a central hub, where it could be redistributed among family and non-family members (Gamble 1999). The redistribution process, at least for earlier hominids, is called "food sharing". But this relatively "warm and fuzzy" term obfuscates the redistribution process. Isaac, who pioneered the notion of communal food consumption, says (1993, p. 535), "The food-sharing model has been widely misunderstood as implying…friendly, cuddly, cooperative human-like hominids. This need not be so." He goes on to suggest food-sharing is really a behavioral system that facilitates exchange. Ofek suggests the term "redistribution" or "exchange" replace the term "food sharing".

including geography and the rate of evolution, may also affect which outcome arises. In particular, we find that speech is less likely to evolve the slower is the evolutionary rate of speech relative to the trading system. This is essentially the *Red King* effect described by Bergstrom and Lachman (2003): the faster player (hominids in this case) loses the evolutionary race, in contrast to the Red Queen effect in which the faster competitor wins. But with more favorable geography, speech and trade may always emerge.

Lieberman (1984) theorizes that multiple evolutionary equilibria did occur as *Homo heidelbergensis* speciated into Neanderthals (*Homo neanderthalensis*).[7] He indicates some SVT changes likely occurred in *Homo heidelbergensis* that allowed for (i) mouth breathing, which increased aerobic efficiency and created a selective value for physical activities, and (ii) proto-speech (primitive speech abilities), but not human speech. At this point there could have been evolutionary gains from further investments in muscular ability (due to more efficient respiration under a non-linguistic morphology) or improved speech capabilities. But the physiologies of these investments are at odds since they involve different uses of the upper respiratory system.

Neanderthals evolved to become more robust and stronger than modern humans (Klein 2003; Trinkhaus and Shipman 1993; Lovejoy and Trinkhaus 1980). Lieberman (1984, 2007) suggests Neanderthals did not evolve a human SVT (though the jury is still out on their vocal skills; Krause et al. 2007).[8] Moreover, Bruner (2004) indicates the parietal development of early modern humans differed from Neanderthals (and earlier hominids). The human parietal developments, which evolved around the same time as the human SVT (or perhaps in response to it), would have enhanced "visiospatial integration, sensory integration, multimodal processing, and social communication" (p. 299). In terms of behavior, there is evidence that Neanderthals likely did not engage in trade to the same extent as modern humans (Tattersall and Schwartz 2000; Tattersall et al. 1998; Kuhn and Stiner 1998; see also Horan et al. 2005). Tattersall (personal communication) suggests that modern socio-economic behaviors are expressions of a generalized underlying capacity that was recently acquired and which Neanderthals did not have.

Our discussion proceeds as if Neanderthals did not have human vocal abilities or trade, though our analysis does not actually require that this was the case. Our model generates multiple equilibria, and it is certainly possible that the conditions for attaining the speech-trade equilibrium were met by Neanderthals. But if Neanderthals did not have these traits, the multiple equilibria arising in our model might help to explain this. Moreover, when viewed alongside Horan et al.'s (2005) theory that Neanderthals went extinct due to inefficiencies caused by their lack of trade relative to humans, our model addresses the natural follow-up

---

[7] Scientific classification of the hominid phylogeny has changed in recent years due to new molecular findings (e.g., Ward and Stringer 1997). Lieberman (1998) refers to *Homo erectus* as the direct ancestor to both Neanderthals and *H. sapiens*, and Ofek (2001) refers to early *H. sapiens* as the immediate predecessor to both. The Smithsonian Institute (2005b) states, "For many years, scientists placed any problematic specimens displaying mixtures of "erectus-like" and "modern" traits into a confusing category: "Archaic" *Homo sapiens*… Recently, it has been proposed to separate these individuals into a distinct species. For this purpose,…the specimen name *Homo heidelbergensis* has seniority." Previously, *H. heidelbergensis* was the generic name given to the first hominids in Europe (Gamble 1999). Although *H. erectus* is now viewed on a different lineage than *H. sapiens* and Neanderthals (Smithsonian 2005a), many of the works cited here still refer to *H. erectus* as the immediate ancestor to both.

[8] There is indirect evidence that Neanderthals, who supported a complex culture (Gamble 1999), had primitive speech (Lieberman 1984). Neanderthal brains were also larger than those of modern humans (Corballis 2004), indicating the potential cognitive capacity for speech (Lieberman 1984, 2007). But (Lieberman 1984, 2007) indicates that a Neanderthal skull could not have supported a human SVT.

question of why some species trade and others do not.[9] In doing so, we also formalize Mellars (2004) speculation that complex language was the key to human success over Neanderthals.

2 Hominid biology

Consider N individuals (or households) of hominids (say *H. heidelbergensis*) living on an area of size K. Following conventional population growth models (McGehee and Armstrong 1977; Dobson 2004), define per capita growth of the population, or net per capita fertility, by

$$\dot{N}/N = G = -d + b(N)F \tag{1}$$

where d is the mortality rate, bF is the birth rate, F represents per capita consumption, and b is a density-dependent birth rate parameter (with $b_N < 0$). Density-dependent growth implicitly accounts for resource availability, which is likely to differ in different geographies (e.g., different latitudes). We return to the geographic issue in the numerical simulation of Sect. 4. Equation 1 indicates the population grows when nutrition is consumed at a rate greater than a subsistence level, defined as $S = d/b$.

2.1 The role of SVT morphology

We now introduce SVT morphology into our model, accounting for the tradeoffs this morphology implies between vocalization and strength.[10] Define the variables v and s to represent the physiological traits of vocalization and strength. We model these two traits as functions of an underlying physiological trait, denoted by the scalar $I \in [0, 1]$. I can be thought of as the degree to which the SVT has evolved towards the human SVT: larger values of I represent a greater degree of vocalization and less strength, while smaller values of I represent less vocal ability and greater strength. Vocalization is given by the continuous function $v(I)$, with $v_I > 0$. Denote strength by $s(I)$, which is decreasing in $I (s_I < 0)$.

These traits enter per capita growth as follows. Mortality is increasing in vocalization ($d_v > 0$) because the associated drop in the larynx leads to a greater risk of choking (Lieberman 1998). The birth rate parameter is decreasing in mass ($b_s < 0$). Larger animals tend to have lower birth rates, as more energy is required to support greater mass, diverting energy flows away from reproductive activity.[11] Subsistence requirements are increasing in muscle mass and in vocalization, so that greater vocalization ($I > 0$) has an ambiguous impact on subsistence requirements.

---

[9] Indirectly, our model is a logical complement to Robson's (2005) social intelligence model and Robson and Kaplan's (2003) theory of coevolutionary investments in intelligence and longevity. Their approaches differ from our model, although we expect these investments could be related in that brain development allows for more sophisticated and clever communication, provided one also has the vocal ability to articulate these superior ideas.

[10] We do not specifically model investments in intelligence via the parietal lobe. These could have occurred along with or as a result of speech and trade coevolution. But many other factors besides these were also likely involved, and it is impractical to attempt to model everything (in our Discussion Section we explore in more detail how the tradeoffs associated with SVT morphology investments likely compared to those of intelligence investments). Robson's (2005) discussion of social intelligence theory suggests that competition for resources among hominids likely led to increased intelligence via the Red Queen effect. This effect could have been heightened once trade became more prevalent.

[11] This is generally true across species, and so it should also be expected to hold as one species (e.g., *H. heidelbergensis*) evolves into another (e.g., *H. sapiens*).

The SVT also affects production and exchange, both of which affect F. Each individual is assumed to produce $Y(s)$ in each period, with stronger individuals producing more ($Y_s > 0$). Also, we describe in the next section that individuals who seek to trade with others incur transactions costs associated with searching for a potential trading partner and communicating about a trade. These costs are denoted $T(v, n)$, where $n = N/K$ represents hominid density. Assume $T_v < 0$ and $T_n < 0$: improved vocal ability reduces communication costs, and a denser population reduces search costs. Here F depends on both v and s (as well as N).

3 Exchange with transactions costs: static and dynamic models

Per capita population growth, G, is influenced by whether individuals choose to participate in exchange/trade. For notational ease, we model trade implicitly through a "reduced-form approach" (see Horan et al. 2005 for an explicit model of paleo-trade). Each hominid is a potential trader. An individual's economic decision is whether to produce and consume only his own output (nutrition), or to specialize and trade with others. An individual who neither specializes nor trades produces and consumes a nutrition level of $Y(s)$. Someone who specializes and trades consumes $Y(s) + Z - T(v, n)$, where Z represents the gross gains from trade relative to the no specialization/no trade scenario.[12] The net gains from trade are endogenous in our model, even if the gross gains are a parameter, because population density enters the model through its attenuating effect on transaction costs. Transactions costs arise because individuals who decide to trade must search for a potential trading partner and then communicate about a trade. For simplicity, assume an individual who incurs these costs finds *one* potential trading partner (the first person he meets) with certainty and attempts to trade with *only* this person; no additional searching occurs if a trade does not occur.[13] An individual who specializes but is unable to find a trading partner consumes $\kappa Y(s) - T$, with $\kappa \in (0, 1)$: while specialization increases the production of nutritional components for which he specializes, a lack of variety in the absence of trade generally decreases aggregate nutritional value.

Whether a trade ultimately occurs depends on whether the potential trading partner is willing to trade. Denote $\sigma \in [0, 1]$ to be the trading strategy adopted by others in the population, i.e., the overall likelihood that a person he meets will trade. A trader consumes nutrition of $Y + Z - T$ with probability (or belief) σ, and he consumes nutrition of $\kappa Y - T$ with probability $1 - \sigma$, so that his expected consumption level is $\sigma(Y + Z - T) + (1 - \sigma)(\kappa Y - T)$.

3.1 Hominid strategies and strategic complementarities in a one-shot game

What proportion of the population will search for a trading partner? In the next section we model the decision to trade as a dynamic process that adjusts via replicator dynamics. But

---

[12] These exogenous gross gains could be modeled to be even greater when more people participate in the market, i.e., $Z = Z(\sigma N)$, with $Z' > 0$. Again adding this feature unnecessarily complicates the model without affecting the primary results. We can also derive our results from the implicit trading using a more notationally complex Ricardian trade model, though again this unnecessarily complicates the model without affecting our primary results.

[13] At the end of the period there is a limited window of opportunity to trade your output (or else it goes bad, or else it gets dark, etc.). Assume two people meet with certainty and this is the only opportunity to trade, but all the results spill over (again with more notational clutter) if instead we assume specialists can potentially meet Q people (as opposed to one).

first we use a one-shot game to illustrate the fundamental concepts underlying the dynamic model.

We assume individuals are homogenous in terms of physiologic traits, but that their behavioral choices may differ. An individual hominid makes choices to maximize net fertility, which in our model is consistent with maximizing nutrition. Each individual i $(i = 1, \ldots, N)$ decides on a trading strategy, denoted by $\rho_i \in [0, 1]$, which represents the probability he searches for a trading partner. His expected consumption is given by

$$F_i = \rho_i(\sigma[Y(s) + Z - T(v, n)] + [1 - \sigma][\kappa Y(s) - T(v, n)]) + (1 - \rho_i)Y(s). \quad (2)$$

In a one-shot game, the optimal choice of the trading probability, ρ_i, is determined by taking the derivative of F_i with respect to ρ_i

$$\frac{\partial F_i}{\partial \rho_i} = \sigma[Z + (1 - \kappa)Y(s)] - T(v, n) - (1 - \kappa)Y(s). \quad (3)$$

The right hand side (RHS) of (3) is the expected net gain from trading. If one decides to trade, he incurs transactions costs and a loss of aggregate nutrition with probability one, but he has a chance to recoup the nutritional losses and earn an additional gain with probability σ.

If the RHS of (3) is negative, then the optimal decision is not to trade: $\rho_i = 0$. In a symmetric Nash equilibrium, $\rho_i = \sigma = 0$ and there is no trade. This outcome always occurs when transactions costs exceed the gains from trade, $T > Z$; and it may occur even when $Z > T$. $Z > T$ is a necessary (but not sufficient) condition for the RHS to be non-negative; the possibility exists for some trading to occur. If the RHS is strictly positive, $\rho_i = 1$ and in a symmetric Nash equilibrium everyone always trades ($\rho_i = \sigma = 1$). In this case, (3) reduces to $Z - T > 0$. If $Z > T$ and the RHS of (3) vanishes, then $\rho_i \in (0, 1)$ and in a symmetric Nash equilibrium we have $\rho_i = \sigma = \hat{\sigma} = (\hat{Y} + T)/(\hat{Y} + Z) < 1$ where $\hat{Y} = (1 - \kappa)Y(s)$. Assuming $Z > T$, the ultimate equilibrium depends on expectations of the individuals. A person wants to mimic others: trade if others trade; no trade if others do not trade; and randomize his trading behavior if others randomize.

Figure 1 illustrates the relation between ρ_i and the relative transactions costs, T/Z. For a given value of T/Z, each individual's expectations about the trading behavior of others must lie on or above the $\hat{\sigma}$ curve for a trading equilibrium to emerge. As the relative transactions costs are increased, a trading equilibrium is less likely; it can only arise if there are sufficient expectations for trading. Finally, no trading occurs when relative transactions costs are sufficiently large, or $T/Z > 1$. Any one of three equilibria could be realized as long as $T/Z < 1$, and so exchange is not guaranteed. We focus on the cases in which $T/Z < 1$, such that all equilibria have a chance to emerge.

The interior equilibrium strategy, which can be written as $\hat{\sigma} = [T/Y + (1 - \kappa)]/[Z/Y + (1 - \kappa)]$, is the ratio of costs to expected benefits. This means the equilibrium trading strategy $\hat{\sigma}$ is reduced when the relative benefits of trading are increased. This result seems counter-intuitive in a static sense, but the intuition becomes clear when dynamic adjustments are considered. In a dynamic context, which we explore in the following section, $\hat{\sigma}$ represents a line of unstable equilibria as indicated by the phase arrows in Fig. 1. If $\sigma > \hat{\sigma}$, the $\sigma = 1$ equilibrium emerges; if $\sigma < \hat{\sigma}$, the $\sigma = 0$ equilibrium arises. Anticipating future results, we note the basin of attraction for trade increases as transaction costs fall (e.g., when population densities are higher or when geographical conditions for supporting human populations are more favorable).

Multiple equilibria may arise when there are strategic complementarities between the individual's strategy and the strategies of others (see Cooper and John 1988; Kiyotaki and Wright 1993; Williamson and Wright 1993; Krugman 1991). The trading strategies of the
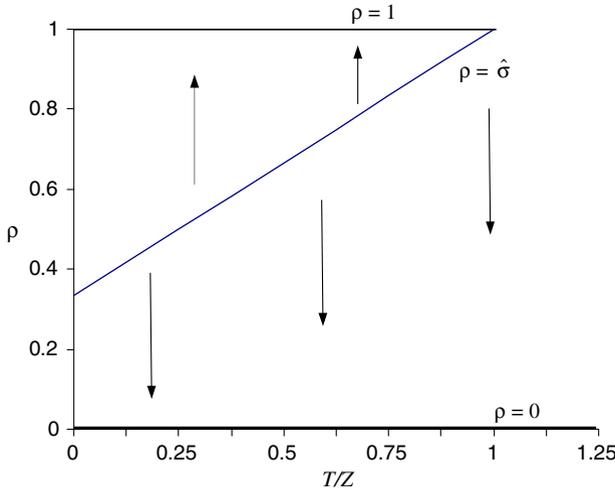
Fig. 1 Interior equilibria of the static model (with phase arrows drawn in to indicate the need to consider dynamics)

individual, ρ, and the others, σ, are *strategic behavioral complements* when the derivative

$$\partial^2 F_i / (\partial \rho_i \partial \sigma) = Z + (1 - \kappa)Y \tag{4}$$

is positive, which it always is: the marginal returns to one person's actions are increased when there is an increase in the others' actions (accordingly, this is a supermodular game; Levin 2003). We adopt the terminology of *strategic behavioral complements*, as opposed to the more standard term *strategic complements*, to distinguish between hominid strategies and evolutionary strategies, which we discuss below.

We now carry the notion of strategic complementarities to the evolutionary scale, with hominid choices depending on the evolution of trait I and vice versa. Consider the impact of increased vocalization strategies on the benefits of trading strategies. Differentiating (3) we see the derivative

$$\partial^2 F_i / (\partial \rho_i \partial I) = -T_v v_I - (1 - \sigma)(1 - \kappa)Y_s s_I \tag{5}$$

is positive: vocalization increases the marginal benefits of trading, and so vocalization is a *strategic coevolutionary complement to trading.*[14]

3.2 Replicator dynamics for the trading strategy

Since the static model provides no intuition into how strategies evolve over time or whether equilibrium strategies are stable, we reformulate the model as a dynamic one. Population dynamics are again given by Eq. 1. Again we assume physiologically homogenous individuals, and we assume a symmetric trading strategy σ. The replicator dynamics for the trading strategy take the form (Rice 2004; Bergstrom and Lachman 2003):

---

[14] In the evolutionary game theory literature, it is common to say that nature does strategize (e.g., Rice 2004), although technically this is inaccurate. Rather a selection process occurs within the natural system that makes it seem *as if* nature chooses a strategy to maximize fitness. If we take this realistic view, it is more appropriate to say vocalization is a joint evolutionary complement to trading (as opposed to a strategic evolutionary complement).

$$\dot{\sigma} = \alpha \left(\sigma [F_T^* - F^*] + (1-\sigma)\delta \right) = \alpha \left(\sigma(1-\sigma)[\partial F/\partial \rho] + (1-\sigma)\delta \right)$$

$$= \alpha \left(\sigma(1-\sigma)[\sigma Z - T(v, n^*) - (1-\sigma)(1-\kappa)Y(s)] + (1-\sigma)\delta \right) \qquad (6)$$

where F_T denotes expected consumption under trade (defined by (2) with $\rho = 1$), F is mean consumption in the population (defined by (2) with $\rho = \sigma$), and $\delta > 0$ is a strategy muta-tion parameter. The parameter $\alpha \in (0, 1)$ indicates the speed of adjustment. Without loss of generality, assume cultural frictions (e.g., induced by traditions, communication difficul-ties, or historic hostilities) slow changes in hominid strategies so these evolve on a slower time scale than does N. We capture this friction by assuming N attains its equilibrium value prior to adjustment: the superscript (*) denotes that N is evaluated at the steady state value N^*(Fenichel 1979 discusses dynamic systems involving fast-time and slow-time variables; see Grimsrud and Huffaker 2006, for a bioeconomic application). While this is not a necessary assumption, it helps us highlight the tradeoffs between trading and vocalization strategies.

Equation 6 indicates individuals increase their trading strategy if doing so leaves them bet-ter off on average. We also assume there is a constant but small rate δ of mutating one's strategy towards trading (see Foster and Young 1990; Rice 2004). This mutation prevents the system from settling at the autarky corner (where $\sigma = 0$): small efforts to trade may be sustained as $\sigma \to 0$. The effect of the mutation term diminishes and ultimately vanishes as $\sigma \to 1$.[15]

Three equilibrium strategies arise from (6) when $Z > T$: (i) a stable *autarkic* equilibrium with nominal trade supported by mutation, $\underline{\sigma} = [(\hat{Y}+T) - \sqrt{(\hat{Y} + T)^2 - 4\delta(\hat{Y} + Z)}]/[2(\hat{Y} + Z)]$, (ii) an *unstable* equilibrium, $\hat{\sigma} = [(\hat{Y} + T) + \sqrt{(\hat{Y} + T)^2 - 4\delta(\hat{Y} + Z)}]/[2(\hat{Y} + Z)]$, and (iii) a stable *trading* equilibrium, $\sigma = 1$. Note $\underline{\sigma} = 0$ and $\hat{\sigma} = (\hat{Y}+T)/(\hat{Y}+Z)$ (the line of interior equilibria in Fig. 1) when $\delta = 0$. These stability properties stem from the strategic behavioral complementarities. With δ small, $\partial F/\partial \rho \approx 0$ when $\sigma = \hat{\sigma}$. As σ increases from this point, $\partial F/\partial \rho > 0$ (since $\partial^2 F/[\partial \rho \partial \sigma] > 0$) and $\dot{\sigma} > 0$. The opposite occurs as σ is reduced from $\sigma = \hat{\sigma}$.

Holding I fixed, the equilibrium ultimately attained depends on the initial trading strat-egy, denoted $\sigma(0) = \sigma_0$. The equilibrium $\underline{\sigma}$ emerges whenever $\sigma_0 < \hat{\sigma}$, and $\sigma = 1$ arises whenever $\sigma_0 > \hat{\sigma}$. The threshold level $\hat{\sigma}$ depends on the population density, n^*, through the impact of density on transaction costs. Specifically, $\hat{\sigma}'(n^*) < 0$; so trade is more likely to take off and establish itself in geographic regions that support denser populations.

We have shown the establishment of trade depends on initial expectations and environ-mental considerations (as some environments initially support greater densities). Trade also depends on SVT physiology, I. Until now we have held I fixed, but over time this can adjust and affect $\dot{\sigma}$. Increases in I will generally increase the marginal benefits of trad-ing, $\partial^2 F/[\partial \rho \partial I] > 0$, (vocalization is a strategic coevolutionary complement to trading), increasing the basin of attraction for the trading equilibrium.

3.3 Adaptive dynamics for the vocalization strategy

Consider the evolution of the human SVT, which we call *nature's vocalization strategy*. Following the convention of the evolutionary game theory literature as applied to quantitative

---

[15] Accordingly, when we introduce coevolution with I in the next section, a small mutation term is gener-ally insufficient to create a *snowball effect* of run-away vocalization. Note we could also include a term for mutation away from trading (e.g., a small mutation rate of γ so that the term $\sigma \gamma$ is subtracted from Eq. 6), but this would complicate the algebra without impacting on the dynamics when σ is small—the focus of our attention since this is the region where there exists the greatest uncertainty about the emergence of trading. The only significant impact of incorporating mutations away from trading would be to prevent the system from equilibrating at the full trade corner, $\sigma = 1$.

genetics, assume nature does strategize (Rice 2004), although technically this is inaccurate. Rather, a selection process occurs within the natural system that makes it seem *as if* nature chooses a strategy to maximize fitness, G. We describe below how this choice of strategy depends on the level of trade. We also define the *level* of trade necessary to affect human speech evolution, and the *lack* of trade necessary to affect muscular evolution.

Mutations in I follow a process of adaptive dynamics (Brown and Vincent 1987; Rice 2004). A strategy for I is an evolutionary stable strategy (ESS) if

$$\frac{\partial G^*}{\partial I} = 0 \text{ and } \frac{\partial^2 G^*}{\partial I^2} < 0 \tag{7}$$

Following Diekmann and Law (1996), Lande (1979), and Krakauer and Jansen (2002), the evolution of the trait I is specified as follows:

$$\dot{I} = \phi\mu N^* \frac{\partial G^*}{\partial I}, \tag{8}$$

where μ is the mutation rate of I, ϕ is the speed of adjustment, and using (2), G^* is given by

$$G^* = -d(v) + b(s, N^*)F^*. \tag{9}$$

The interpretation of parameter ϕ is analogous to that of the adjustment parameter α introduced in (6); but if ϕ is sufficiently greater or smaller than α, the trading and mutation processes occur at different time scales. If ϕ is very small relative to α, trading dynamics are fast and mutation dynamics are slow; here the superscript ($*$) in (8) and (9) applies to both the trading strategy σ and the population N. If ϕ and α have similar magnitudes, the superscript ($*$) only applies to N. If ϕ is very large relative to α, trading dynamics are slow and mutation dynamics are fast. For now, assume ϕ and α are of similar magnitudes; in our numerical simulation later in the paper we explore how changes in the relative magnitudes affect the dynamics.

From (8), trait I attains a steady state when $\partial G^*/\partial I = 0$, or when proportional net growth (G^*) is maximized as is required by (7). Otherwise, there are increases in vocalization (strength) when the marginal product of the trait on fitness is positive (negative). Interestingly, investments in vocalization depend on the trading strategy. To see this, differentiate (9) to obtain:

$$\frac{\partial^2 G^*}{\partial I \partial \sigma} = b_s s_I \left[\frac{\partial F^*}{\partial \rho} + \sigma \frac{\partial^2 F^*}{\partial \rho \partial \sigma} \right] + b \left[\frac{\partial^2 F^*}{\partial I \partial \rho} + \sigma \frac{\partial^3 F^*}{\partial I \partial \sigma \partial \rho} \right] + \frac{\partial^2 G^*}{\partial I \partial N^*} \frac{\partial N^*}{\partial \sigma} \tag{10}$$

where $\partial F^*/\partial \rho$ is as defined in (3), $\partial^2 F^*/(\partial \rho \partial \sigma)$ is as defined in (4), and $\partial^2 F^*/(\partial I \partial \rho)$ is as defined in (5).[16] The first RHS term is of the same sign as the expression in brackets, which consists of the marginal impact of an individual's strategy on his own consumption ($\partial F^*/\partial \rho$, which may be positive or negative) plus the weighted (by σ) strategic behavioral complementarity effect of others' strategies on his consumption ($\sigma \partial^2 F^*/(\partial \rho \partial \sigma)$, which is positive). Assuming δ is sufficiently small, $\partial F^*/\partial \rho \geq 0$ when $\sigma \geq \hat{\sigma}$, in which case the first bracketed term is positive; the term can be negative if $\partial F^*/\partial \rho$ is sufficiently negative.

The second RHS term in (10) has the same sign as the expression in brackets, which consists of the strategic coevolutionary complementarity effect of vocalization on trade

---

[16] The derivatives in expression (10) are taken prior to imposing the equilibrium condition $\rho = \sigma$. If the equilibrium condition was imposed first, then the expression in the first set of brackets would be written $\partial F^*/\partial \sigma$ and the expression in the second set of brackets would be written $\partial^2 F^*/(\partial \sigma \partial I)$. We have written the derivatives in the current form to highlight the role of strategic complementarities.

$(\partial^2 F^*/(\partial\rho\partial I)$, which is positive) plus the weighted marginal impact of an individual's strategy on this complementarity $(\sigma\partial^3 F^*/(\partial I\partial\sigma\partial\rho) = \sigma(1-\kappa)Y_s s_I < 0)$. When σ is sufficiently small (i.e., trade has only just emerged), then $\sigma(1-\kappa)Y_s s_I$ is small and the overall bracketed term is positive. Even when $\sigma = 1$, the overall term is positive if increased vocalization has greater marginal impacts on transactions costs than on output weighted by $(1-\kappa)$. Finally, the third RHS term in (10) reflects the impact of an increase in σ on the equilibrium population level N^*, which may be positive or negative but is presumably small if the net gains from trade are small.

Assuming the final two terms in (10) are positive or negligible, trading is a strategic coevolutionary complement to vocalization $(\partial^2 G^*/(\partial I\partial\sigma) > 0)$, provided the individual's marginal benefits of trading $(\partial F^*/\partial\rho)$ are positive or not too negative. In this case, an increase (decrease) in the trading strategy reinforces the incentives to increase (decrease) the vocalization strategy and vice versa. Sufficient levels of trade could therefore be a necessary condition for evolution into modern humans, while a sufficient lack of trade could lead to Neanderthal-style evolution.

In contrast, if the individual's marginal benefits of trading are sufficiently negative, $\partial^2 G^*/(\partial I\partial\sigma)$ may be negative. Trading would be a strategic coevolutionary substitute to vocalization in this case (increasing the trading strategy reduces the benefits of vocalization), although vocalization remains a strategic coevolutionary complement to trading. The net result is that a hominid society could become trapped in an equilibrium involving low levels of trade and vocalization, as the investment incentives are not self-reinforcing.

4 Simulation

We now use a simulation (specified in Table 1) to show the two ways in which the trade-vocalization equilibrium can emerge in our model. The equilibrium can emerge (1) due to favorable initial conditions in the presence of multiple equilibria caused by adverse geographic conditions supporting small population densities, or (2) as a globally optimal solution such as when geography supports "sufficiently" large population densities. In the case of multiple equilibria, the key influence on the necessary initial conditions is the relative speed of adjustment of speech and trade. When the evolutionary speeds are asymmetric (i.e., $\alpha \neq \phi$), the Red King effect kicks in for better or worse, depending on initial conditions.

4.1 Scenario A: poor geography and multiple equilibria

We consider two scenarios in which the local geography does not support dense populations. Scenario A, illustrated in Fig. 2, represents a situation in which both strategies, ρ and σ, evolve along an identical time scale, i.e., $\phi = \alpha$. Three equilibria emerge: x is a locally stable, autarkic-strength equilibrium; y is a saddle-point equilibrium; and z is a locally stable, vocal-trade equilibrium. The saddle path, denoted by S, is the only trajectory that leads to y. But the probability the system is initially on that path is essentially zero. Therefore, y is effectively unstable and S separates the phase plane into two basins of attraction: one for x and one for z.

All points above S are on trajectories leading to z. The reason is that, in this region of the phase plane, vocalization and trade are bi-lateral strategic coevolutionary complements of sufficient magnitude as to overcome the hominid coordination failures. Vocalization and trading strategies strongly reinforce each other in this region, making z locally stable. Contrast this outcome with what happens in the absence of coevolution. Suppose society is initially at

Table 1 Simulation model equations and parameter values

Biological process	Equations	Parameter values	Notes
Vocalization	$v = v_0(1 + v_1 I)$	$v_0 = 0.5$ $v_1 = 1$	
Strength	$s = s_0(1 - s_1 I)$	$s_0 = 1$ $s_1 = 0.75$	
Mortality	$d = d_0(1 + d_1 v)$	$d_0 = 0.07$ $d_1 = 0.15$	
Birth	$b = b_0(1 - b_1 s) - b_2 N$	$b_0 = 0.35$ $b_1 = 1.2$ $b_2 = (b_0 y(s(0.5)) - d_0)/$ $(y_0 k)$	b_2 is the density-dependent fertility term, calibrated at an initial equilibrium (i.e., where $G = 0$) with mid-range level of strength, $I = 0.5$, so that initial output at this point is $y(s(0.5))$,
		$k = 0.7$ for Scenarios A and B; $k = 1.0$ for Scenario C	k is the carrying capacity, expressed as a density
Output	$Y = Y_0 s$	$Y_0 = 0.5$	
Transactions cost	$T = T_0(1 - v)/(N/K)$	$T_0 = 1$ $K = 1$	
Gains from trade		$Z = 0.75$	
Mutation rate		$\mu = 1$	The impact of μ cannot be separated from that of ϕ, as the effective speed of adjustment is $\mu\phi$
Speed of adjustment		$\alpha = 0.0002$ in Scenario A; $\alpha = 0.02$ in Scenario B $\phi = 0.0002$	Our discussion of the speed of adjustment is centered around the separatrices, which do not exist in Scenario C
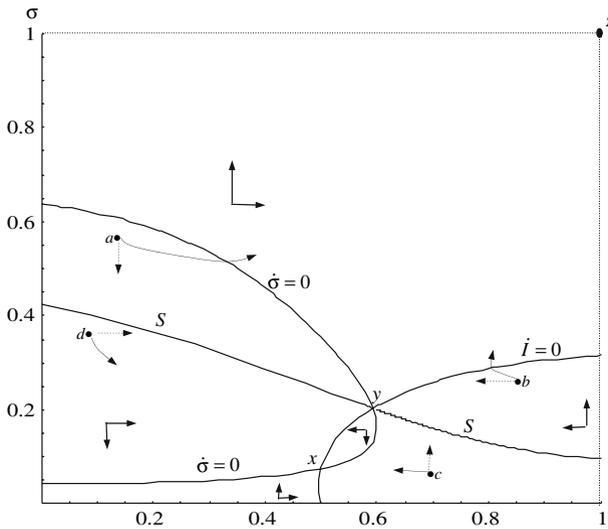
Fig. 2 Scenario A—symmetric evolutionary dynamics

point a. If I were fixed, the trade trajectory would follow the dotted trajectory to a low trade equilibrium on the $\dot{\sigma} = 0$ isocline. Alternatively, suppose society started at point b, and let σ be fixed. Now I would follow the dotted trajectory until it equilibrates at a lower level of vocalization on the $\dot{I} = 0$ isocline. This implies the introduction of coevolution allows both trading and vocalization to evolve from either of these points along the solid trajectories to equilibrium z. Coevolution generally improves the outcome.

In contrast, all initial points below S follow trajectories to x. The local stability of x in this region of the phase plane occurs because trade and vocalization are not strong strategic coevolutionary complements to one another. As a result, the reinforcing incentives are small. The increased trading incentives that come from increased vocalization are not enough to overcome the hominid coordination failures. Moreover, if people trade less then incentives for nature to invest in vocalization decline. The hominid society gets stuck in an autarkic-strength equilibrium. We again contrast this outcome against one without coevolution. Suppose society is initially at point c. If I were fixed, the trade trajectory would follow the dotted trajectory to $\sigma = 1$. Alternatively, suppose the initial point is d and σ is fixed. Now I would follow the dotted trajectory to $I = 1$. Coevolution here results in the solid trajectories leading to equilibrium x; coevolution generally worsens the outcome.

Hominids (as a group) prefer equilibrium z to x, whereas nature is indifferent between these equilibria since $G^* = 0$ holds at each equilibria. While the basin of attraction for z is large, getting there requires sufficient hominid coordination—a sufficiently large σ_0. The required value of σ_0 diminishes as vocalization (I_0) is increased, but some positive value of σ_0 is always required to enter the basin of attraction for equilibrium z.

4.2 Scenario B: poor geography, multiple equilibria, and the Red King effect

Scenario B (Fig. 3) captures the same situation as Scenario A except now we assume asymmetric coevolution, in which the trading strategy evolves faster, i.e., $\phi < \alpha$. The effect is to rotate S clockwise, *shrinking* the basin of attraction for the hominids' preferred equilibrium outcome, z.[17] This is the Red King effect described by Bergstrom and Lachman (2003)—the basin of attraction for the slower (faster) player's preferred outcome expands (shrinks). Intuitively, the smaller basin of attraction results because the faster player's effective "bargaining power" is diminished.[18] In our problem, less of the isosector located to the left of the $\dot{\sigma} = 0$ isocline remains in the basin of attraction for z. The trading strategy is diminishing over time in this region. When this decline occurs relatively quickly it increases the downward pull toward equilibrium x—the greater speed enhances the downward effect caused by coordination failures, overwhelming the counteracting upward effect caused by the complementary

---

[17] In the limit as $\phi \to 0$ (vocalization dynamics are extremely slow relative to trade dynamics), vertical movements seem to occur instantaneously relative to horizontal movements. In this case, the northwest separatrix S is defined by the northwest portion of the $\dot{\sigma} = 0$ isocline, and the southeast separatrix S becomes vertical and tangent to the $\dot{\sigma} = 0$ isocline at the point where the isocline bends backwards. Finally, note that S rotates counterclockwise relative to Fig. 1 when hominid trade evolves more slowly than nature ($\alpha < \phi$). We do not investigate this case formally, but the results and intuition will be the opposite of that explained here.

[18] The Red King effect emerges when two players are involved in a mutualism (a cooperative endeavor in which both players stand to gain) and one player's strategy evolves slower than another's. The intuition offered by Bergstrom and Lachman (2003) for the Red King effect is as follows. Each player involved in the mutualism prefers to be selfish, but as there are multiple equilibria the process of getting to one's preferred outcome is akin to bargaining for that outcome. In Schelling's (1960) early discussion of bargaining, a player stands to gain if his hands are tied by constraints originating from within his clan. This is the so-called Schelling conjecture, in which national negotiators can bargain harder in international agreements when they have domestic constraints they must satisfy. In evolutionary game theory models, a slower speed of adjustment effectively ties one's hands.
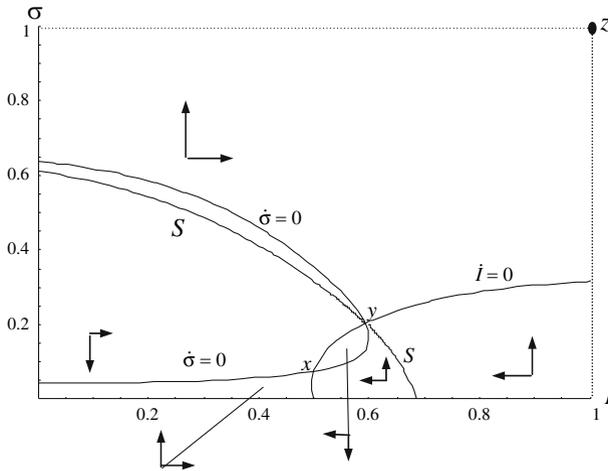
Fig. 3 Scenario B—asymmetric evolutionary dynamics

impact of greater vocalization. That is, the relatively larger value of α effectively reduces the strategic coevolutionary complementarities of vocalization on trade. Moreover, the faster reduction in trade reduces the incentives for vocalization. Together, these forces expand the basin of attraction for x.

That the faster player's preferred basin of attraction is diminished leads Bergstrom and Lachman to conclude that the fastest player loses a Red King game—the opposite of the antagonistic coevolutionary Red Queen game in which the fastest player wins. But they recognize that this global result does not necessarily hold for each smaller region of the phase plane. While reducing the relative value of ϕ in our model reduces the overall basin of attraction for z, portions of this basin actually expand in regions where there has already been sufficient speech (or proto-speech) development—specifically, to the right of the right-hand arm of the saddle path S. Initial conditions matter. Since sufficient vocal abilities exist already in this region, less initial trade (possibly zero) is required to move to the vocalization-trade equilibrium. Moreover, the faster player wins in this region, similar to the Red Queen effect, but for a different reason. The intuition is the reverse of what we discussed earlier. The larger basin of attraction occurs in the region where the trading strategy is increasing over time (i.e., $\dot{\sigma} > 0$). Here, the relatively greater speed at which the trading strategy increases intensifies the upward pull acting on both strategies, effectively increasing the strategic coevolutionary complementarities and expanding the basin of attraction for z.

To sum up, when multiple equilibria exist and if trade evolves fast relative to vocalization, then (i) more initial trade is needed to get the coevolutionary process rolling towards z, at least when there is low initial vocalization, and (ii) less initial trade is needed to move the system towards z when there is a sufficient initial level of vocalization (or proto-speech). One can only speculate as to what the exact initial conditions were, but we know that some degree of proto-speech did exist. If this level were sufficient, trade could take hold even if initial rates were low or zero. But what level of proto-speech is sufficient? The answer depends on the underlying biological and environmental parameters. Geography could have mattered; better geographical conditions will shift the isoclines in ways that expand the basin of attraction for z. We now turn to a scenario in which geographical conditions are improved so much that the trade-vocalization equilibrium is globally stable.
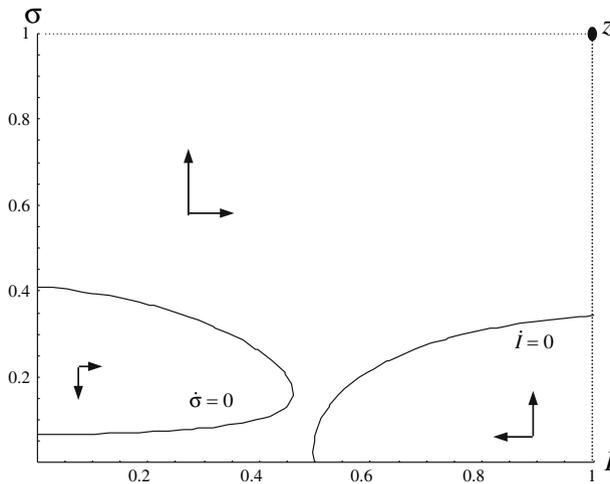
Fig. 4 Scenario C—Better Geographic conditions

4.3 Scenario C: favorable geography and a globally stable equilibrium

Now consider a geographic region that supports a denser population. The larger densities reduce transactions costs of trading relative to less favorable geographies, increasing individuals' marginal benefits of trading. For lower levels of vocalization, the autarkic equilibrium strategy, $\underline{\sigma}$, is increased and the unstable equilibrium strategy, $\hat{\sigma}$, is reduced. For greater levels of vocalization, both of these equilibria disappear (the discriminant of $\underline{\sigma}$ and $\hat{\sigma}$ becomes negative) so that only the trading equilibrium $\sigma = 1$ remains.

Scenario C (Fig. 4) represents a geographic region that exhibits a 20% larger population carrying capacity K (all other parameter values are identical to those in Scenario A). The larger marginal benefits of trade shift the $\dot{\sigma} = 0$ isocline to the left. These larger marginal benefits also increase the marginal benefits of vocalization, though the effect on the $\dot{I} = 0$ isocline is small. The net result is that a bifurcation results relative to Scenario A, as the isoclines no longer intersect and the vocalization-trade equilibrium z becomes globally stable. Given our parameters, the bifurcation occurs with about a 25 % increase in K from the baseline. While this value is simulated, it does suggest the follow pattern—if either environmental or cultural conditions or both improve to support greater hominid densities, a threshold is crossed such that speech via trade emerges as a stable global equilibrium.

4.4 Implications for divergent speciation

The differences between Scenarios A & B and A & C may help explain the distinct evolutionary outcomes of Neanderthals (strong, non-traders with limited vocal abilities) versus humans (weaker, traders with advanced vocal abilities). Earlier we established that lowering the transaction costs of trade increases the basin of attraction for the speech-trade equilibrium. This is confirmed in the numerical analysis—we find for some parameter values the speech-trade outcome is inevitable. Trade could have emerged for more favorable initial expectations about trading, or it could have emerged in regions where cultural changes (i.e., adoption of trade) occurred more rapidly such that initial expectations did not matter as much (as in Fig.

3). In Scenario C, however, initial expectations do not matter—only environmental conditions as created by the geographic region matter. *Better* geographic conditions support greater human densities, and as a consequence, the system settles at *z*.

Our model predicts the speech-trade equilibrium of modern humans is more likely to be the outcome for species with high population densities (due to favorable geography). Were densities of the direct ancestors of *Homo sapiens* in Africa actually greater than densities of the direct ancestors of European Neanderthals? While knowledge about prehistoric population densities is scant, there is reason to believe this was the case. Humans evolved in Africa, in a relatively warm and stable climate. In contrast, Neanderthals lived under more brutal circumstances—making a living in Eurasia in the grip of an ice age. Binford (2001), for instance, estimated populations of hunter–gatherers by ecological zone during the Mesolithic (admittedly a more recent period), based on gradients of diminishing prey availability and growing season as one moves away from the tropics, and the pattern of his findings is consistent with the notion that hominid densities would have been greater in Africa. Keeley (1988) also finds that population density in hunter–gatherer societies is negatively correlated with latitude. Firming up the linkage between speech and trade, Keeley notes only societies with high population densities develop complicated culture characterized by, among other things, "standard valuables or currency" which would enable trade to operate efficiently (p. 405).[19]

We close this section with two speculative observations. First, it is conceivable that our analysis of the *why* of the origin of speech also sheds light on the *when* question. If the co-evolutionary process involving trade is responsible for the development of speech, it is possible the emergence of speech coincides with (or shortly follows) a sudden increase in pre-human population densities in Paleolithic times. This would not be unprecedented. For example, Stiner et al. (1999, p. 190) conclude "human population densities increased abruptly during the late Middle Paleolithic and again during the Upper and Epi-Paleolithic periods." The model predicts that such growth pulses are conducive to the evolution of speech—population shocks, induced by climate or custom, could cast the system into the basin of attraction of the speech-trade equilibrium.

Second, once trade and vocalization take sufficient hold of the population so the system begins to close in on *z*, then migration into regions that support smaller population densities (e.g., represented by a shift in the isoclines that alters the system from Scenario C [Figure 4] to Scenario A [Fig. 2]) will not send the population back to equilibrium *x*. This raises the possibility that humans crossed into Europe without losing any of these abilities.

5 Discussion

The origin of human speech is one prominent *paleo puzzle* in modern science. Our proposed explanation is based on the idea of *strategic complementarities* that result in multiple stable development trajectories—one path gives rise to modern humans, the other gives rise to a muscular species with limited speech capacities (perhaps Neanderthals). Key elements in our model are (i) one's decision to trade affects the returns to similar decisions taken by other agents, (ii) transaction costs are influenced by vocalization abilities, and (iii) trading decisions affect the morphological evolution of humans—favoring investments in muscles or vocalization. Taken together these three elements create conditions favorable for speech

---

[19] Keeley (1998) uses the term "complex" in reference to the particular type of culture being described here, but for consistency we stick with Gamble's (1999) use of the term "complicated". Gamble uses "complicated" to describe human cultures and "complex" to describe the less-sophisticated Neanderthal cultures.

to emerge as a complement to trading, which served humanity by increasing our fertility and as a consequence, population.

Why has the evolution of human speech been so puzzling? One difficulty stems from the fact that human speech relies on neural mechanisms that in present day humans involve parts of the brain that also play a part in motor control, emotion and higher cognition. The apparent result of speech and neural abilities being so intertwined creates two identification problems. First, if an activity was suggested to be an evolutionary driver, how do we know it would have led to developments in speech as opposed to neural abilities (or both)? That is, how do we know the variable I in our model relates to speech and not intelligence? Second, the contribution of human speech to biological fitness would seem to entail most anything anyone can think about, resulting in an identification problem for the causative activity: it is a challenge to relate the evolution of language to any isolated activity like trade.

But focusing on speech and neural abilities confuses matters. While it is the case that higher-order brain function is required for speech, not all higher-order brain functions rely on speech. Many neural investments came earlier, presumably in conjunction with tool-making, and would have supported primitive speech (see the citations in footnote 2). Other neural investments were made at or around the time of speech evolution (footnote 10). But those investments alone would not have directly improved speech capabilities beyond proto-speech because these capabilities were morphologically-limited. The morphology of speech (i.e., physical changes in the SVT) was costly, and its only use is for speech. This morphology most likely evolved specifically for speech, with the benefits stemming from the primary function of speech—improved communication.

The co-evolutionary problem is therefore an economic problem. It is from this perspective that we address the identification problems. The first identification problem is mitigated by focusing on the morphology of speech and the associated costs (unlike prior research). We define the tradeoffs associated with I in accordance with those directly associated with SVT morphology, as this was the limiting feature of speech. A broader interpretation of our model could be that I represents both morphology and intelligence, provided the net costs and net benefits of these two investments are consistent with our specification. Some costs will be the same for the human SVT and intelligence, but some key differences exist. We have modeled greater speech capabilities to result in greater human mortality, less power and consequently less productivity, and greater birth rates. Robson and Kaplan (2003) argue the opposite relationship holds for intelligence and mortality: greater intelligence coevolved with less mortality because longevity is required to yield a return on the knowledge investment that comes from greater intelligence. Similarly, an opposite relation may arise between intelligence and productivity. While brainpower might divert energy from muscles, Robson (2005) notes that intelligence is also needed for effective hunting. If we were modeling brainpower investments the effect would in all likelihood be ambiguous, with a possible productivity advantage going to smarter people—the opposite of what occurs in our model. There may also be an opposite relation between intelligence and birth rates, as energy required for increased brainpower may be diverted from reproduction. Finally, on the benefit side, with more speech one should have reduced transactions costs of trade. Would greater intelligence reduce transactions costs? Not necessarily, if the trader could not speak—again the end result is ambiguous. Still, we admit it is difficult to separate out all the processes that drove the evolution of speech and brain function. Trade-speech co-evolution could have certainly played a role in neural investments, but so could have many other things.

The second identification problem is assuaged by recognizing that the driving co-evolutionary activity (or activities) must be associated with communication. This is not a new insight, as nearly all prior research into this problem focuses on one activity associated

with communication. We focus on trade, which has previously not been formally modeled within this context, and examine this as the trade-off between morphology costs and benefits. It is possible a single activity such as trade could have gotten the ball rolling, and after that there would have been increased benefits from improved communication in a wide range of activities, e.g., mating, hunting.

Another difficulty in identifying the behavior is a lack of behavioral evidence in the archaeological record during this time period. This is a problem for all behavioral explanations, but particularly for trade which is a uniquely human activity that does not arise to a similar extent in other primate systems. Direct evidence of what many people think of as human trade currently goes back about 150,000 years (though evidence of modern human behaviors prior to this time continues to be uncovered; McBrearty and Stringer 2007). Ofek (2001) makes the case that earlier forms of trade have not received adequate attention. Both Ofek (2001) and Isaac (1983) argue that food-sharing, a known practice among even the earliest *Homo* lineages, is a basic process of resource redistribution—one of the earliest forms of market exchange. Ofek suggests some degree of specialization and exchange came to be involved in tool-making and other activities, even among early hominids. While we do not specify the exact nature of exchange in our model, implicitly our exchange environment can be rudimentary and fall somewhere on the spectrum between food-sharing/tool-making and more widespread trade.

Our model frames *one* potential path that this pre-history/pre-evidence co-evolutionary world might have taken—one that accounts for the interaction of biological and economic institutions. Our model addresses the potential co-evolutionary tradeoffs associated with a particular cooperative behavior, trade, and nature's investments in the morphology of speech. Prior theories have also addressed co-evolutionary linkages between speech and behavior (including trade), though most of these theories do not formalize the evolutionary dynamics as we do. In particular, prior work involving trade has not involved a formal model. We have tried to shift the focus to these issues.

At the present time, the complete validity of this world is beyond empirical verification from field data. And obviously simulation models are never useful as "proof" of the occurrence of a certain historic trajectory—they can only be used to evaluate whether certain trajectories are feasible, and whether model outcomes are consistent with stylized facts as currently understood. In this sense the contribution of social scientists to "paleo-puzzles" is complementary to the evidence-driven approaches in the natural sciences and archaeology.

Nevertheless, we believe the model might also be useful to advance the search for new evidence. First, outcomes may help frame where and how to uncover more evidence on the interaction of social and natural systems, e.g., trade, coordination, specialization. Second, the results suggest there is scope for a series of interesting experiments in the laboratory. Lab experiments studying the use of a novel language within coordination games are becoming more common-place (see, e.g., Galantucci 2005; Selten and Warglien 2007). But for the most part these experiments assume all players are capable of producing language and they need to coordinate on a code; the more complicated the code, the more costly. What we are proposing instead is a set of experiments in which the evolution of language can be studied in which the existence of exchange institutions and degree of trade are the treatment variables. While we would not of course be able to simulate biological evolution, we could capture the rates of language development within and outside exchange institutions. That said, researchers would still be left with the challenge of working backwards to connect language development with speech morphology, but with more information about behavior within institutions to guide their search.

References

Baker, M. (2003). An equilibrium conflict model of land tenure in hunter-gatherer societies. *Journal of Political Economy*, *111*, 124–173.

Bergstrom, C. T., & Lachman, M. (2003). The Red King effect: When the slowest runner wins the coevolutionary race. *Proceedings of the National Academy of Sciences, 100*, 593–598.

Binford, L. R. (2001). *Constructing frames of reference*. Berkeley: University of California Press.

Bowles, S., Choi, J.-K., & Hopfensitz, A. (2003). The co-evolution of individual behaviors and social institutions. *Journal of Theoretical Biology, 223*, 135–147.

Brander, J. A., & Taylor, M. S. (1998). The simple economics of Easter Island: A Ricardo-Malthus model of renewable resource use. *American Economic Review, 88*, 119–138.

Brown, J. S., & Vincent, T. L. (1987). Coevolution as an evolutionary game. *Evolution, 41*, 66–79.

Bruner, E. (2004). Geometric morphometrics and paleoneurology: Brain shape evolution in the Genus *Homo*. *Journal of Human Evolution, 47*, 279–303.

Bulte, E. H., Horan, R. D., Shogren, J. F. (2006). The economics of pleistocene megafauna extinction: Early humans and the overkill hypothesis. *Journal of Economic Behavior and Organization 59*, 297–323.

Cooper, R., & John, A. (1988). Coordinating coordination failures in Keynesian models. *The Quarterly Journal of Economics, CIII*, 441–463.

Corballis, M. C. (2002). *From hand to mouth: The origins of language*. Princeton: Princeton University Press.

Deacon, T. W. (1997). *The symbolic species*. New York: Norton.

Dessalles, J.-L. (2000). Language and hominid politics. In C. Knight, M. Studdert-Kennedy, & J. R. Hurford (Eds.), *The evolutionary emergence of language: Social function and the origins of linguistic form*. Cambridge: Cambridge University Press.

Diamond, P. (1982). Aggregate demand management in search equilibrium. *Journal of Political Economy, 90*, 881–894.

Diekmann, O., & Law, R. (1996). The dynamical theory of coevolution: A derivation from stochastic ecological process. *Journal of Mathematical Biology, 34*, 579–612.

Dobson, A. (2004). Population dynamics of pathogens with multiple host species. *The American Naturalist, 164*, S64–S78.

Dunbar, R. I. M. (1996). *Grooming gossip and the evolution of language*. London: Faber and Faber.

Fenichel, N. (1979). Geometric singular perturbation theory for orindary differential equations. *Journal of Differential Equations, 31*, 53–98.

Foster, D., & Young, P. (1990). Stochastic evolutionary game dynamics. *Journal of Theoretical Biology, 38*, 219–232.

Galantucci, B. (2005). An experimental study of the emergence of human communication systems. *Cognitive Science, 29*, 737–767.

Galor, O., & Moav, O. (2002). Natural selection and the origin of economic growth. *Quarterly Journal of Economics, 117*, 1133–1191.

Gamble, C. (1999). *The paleolithic societies of Europe*. Cambridge: Cambridge University Press.

Gifford, A. Jr. (2002). The evolution of the social contract. *Constitutional Political Economy, 13*, 361–379.

Grimsrud, K. M., & Huffaker, R. (2006). Solving multidimensional bioeconomic problems with singular-perturbation reduction models: Application to managing pest resistence to pesticidal crops. *Journal of Environmental Economics and Management, 51*, 336–353.

Hansson, I., & Stuart, C. (1990). Malthusian selection of preferences. *American Economic Review, 80*, 529–544.

Holden, C. (1998). No last word on human origins. *Science, 282*, 1455.

Horan, R. D., Bulte, E. H., & Shogren, J. F. (2005). How trade saved humanity from biological exclusion: An economic theory of Neanderthal extinction. *Journal of Economic Behavior and Organization, 58*, 1–29.

Isaac, G. (1983). Aspects of human evolution. In D. S. Bendall (Ed.), *Essays on evolution: A Darwin century volume*. Cambridge: Cambridge University Press.

Keeley, L. H. (1988). Hunter–gatherer economic complexity and population pressure: A cross-cultural analysis. *Journal of Anthropological Archaeology, 7*, 373–411.

Kiyotaki, N., & Wright, R. (1993). A search-theoretic approach to monetary economics. *American Economic Review, 83*, 63–77.

Klein, R. G. (2003). Whither the Neanderthals? *Science, 299*, 1525–1526.

Knight, C. (1998). Ritual/speech coevolution: A solution to the problem of deception. In J. R. Hurford, M. Studdert-Kennedy, & C. Knight (Eds.), *Approaches to the evolution of language: Social and cognitive bases.* Cambridge: Cambridge University Press.

Knight, C., Studdert-Kennedy, M., & Hurford, J. R. (2000). Language: A Darwinian adaptation? In C. Knight, M. Studdert-Kennedy, & J. R. Hurford (Eds.), *The evolutionary emergence of language: Social function and the origins of linguistic form.* Cambridge: Cambridge University Press.

Krakauer, D. C., & Jansen, V. A. A. (2002). Red Queen dynamics of protein translation. *Journal of Theoretical Biology, 218*, 97–109.

Krause, J., Lalueza-Fox, C., Orlando, L., Enard, W., Green, R. E., Burbano, H. A., Hublin, J.-J., Hänni, C., Fortea, J., de la Rasilla, M., Bertranpetit, J., Rosas, A., & Pääbo, S. (2007). The derived *FOXP2* variant of modern humans was shared with Neandertals. *Current Biology, 17*(Nov. 6), 1–5.

Krugman, P. (1991). History versus expectations. *The Quarterly Journal of Economics,106*, 651–667.

Kuhn, S. L., & Stiner, M. C. (1998). Middle palaeolithic 'Creativity': Reflections on an oxymoron, Chap. 9. In S. Mithen (Ed.), *Creativity in human evolution and prehistory* (pp. 143–164). London: Routledge.

Lagerlöf, N. P. (2005). *Slavery and other property rights*. York University Mimeo.

Lande, R. (1979). Quantitative genetic-analysis of multivariate evolution, applied to brain–body size allometry. *Evolution, 33*, 402–416.

Levin, J. (2003). Supermodular games. Lecture notes, Department of Economics, Stanford University.

Lieberman, P. (1984). *The biology and evolution of language*. Cambridge: Harvard University Press.

Lieberman, P. (1998). *Eve spoke*. New York: W.W. Norton and Co.

Lieberman, P. Current views on Neanderthal speech capabilities: A reply to Boe et al. (2007). *Journal of Phonetics, 35*, 552–563.

Livingstone, D., & Fyfe, C. (2000). Modelling language-physiology coevolution. In C. Knight, M. Studdert-Kennedy, & J. R. Hurford (Eds.), *The evolutionary emergence of language: Social function and the origins of linguistic form*. Cambridge: Cambridge University Press.

Lovejoy, C. O., & Trinkhaus, E. (1980). Strength and robusticity of the Neandertal tibia. *American Journal of Physical Anthropology, 53*, 465–470.

Marceau, N., & Myers, G. M. (2006). On the early Holocene: Foraging to early agriculture. *Economic Journal, 116*, 751–772.

McBrearty, S., & Stringer, C. (2007). The coast in colour.*Nature, 449*, 793–794.

McGehee, R., & Armstrong, R. (1977). Some mathematical problems concerning the ecological principle of competitive exclusion. *Journal of Differential Equations, 23*, 30–52.

Mellars, P. (2004). Neanderthals and the modern human colonization of Europe. *Nature, 432*, 461–465.

Noble, J. (2000). Cooperation, competition and the evolution of prelinguistic communication. In C. Knight, M. Studdert-Kennedy, & J. R. Hurford (Eds.), *The evolutionary emergence of language: social function and the origins of linguistic form*. Cambridge: Cambridge University Press.

Ofek, H. (2001). *Second nature: Economic origins of human evolution*. Cambridge: Cambridge University Press.

Olsson, O., & Hibbs, D. (2005). Biogeography and long-run economic development. *European Economic Review, 49*, 909–938.

Paillard, J. (2000). The neurobiological roots of rational thinking. In H. Cruse, J. Dean, & H. Ritter (Eds.), *Prerational intelligence: Adaptive behavior and intelligent systems without symbols and logic*. Dordrecht: Kluwer.

Pinker, S., & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences, 13*, 707–784.

Rice, S. H. (2004). *Evolutionary theory: Mathematical and conceptual foundations*. Sunderland, MD: Sinauer.

Robson, A. J. (2005). Complex evolutionary systems and the Red Queen. *The Economic Journal, 115*, F211–F224.

Robson, A. J. & Kaplan, H. S. (2003). The evolution of human life expectancy and intelligence in hunter–gatherer economies. *American Economic Review, 93*, 150–169.

Schelling, T. (1960). *The strategy of conflict*. Cambridge: Harvard University Press.

Selten, R. & Warglien, M. (2007). The emergence of simple languages in an experimental coordination game. *Proceedings of the National Academy of Sciences, 104*, 7361–7366.

Smith, V. L. (1975). The primitive hunter culture, Pleistocene extinction and the rise of agriculture. *Journal of Political Economy, 83*, 717–755.

Smithsonian Institution. (2005a). *Early human phylogeny*. http://www.mnh.si.edu/anthro/humanorigins/ha/a_tree.html (accessed August 30).

Smithsonian Institution. (2005b). *Homo heidelbergensis*. http://www.mnh.si.edu/anthro/humanorigins/ha/heid.htm (accessed August 30).

Stiner, M., Munro, N., Surovell, T., Tchernov, E., & Bar-Yosef, O. (1999). Paleolithic population growth pulses evidenced by small animal exploitation. *Science, 283*, 190–194.

Tattersall, I., Delson, E. & van Couvering, J. (1988). *Encyclopedia of human evolution and prehistory*. New York: Garland Publishing.

Tattersall, I., & Schwartz, J. H. (2000). *Extinct humans*. New York: Westview Press.

Trinkhaus, E., & Shipman, P. (1993). *The Neandertals: Changing the image of mankind*. New York: Knopf.

Ward, R., & Stringer, C. (1997). A molecular handle on the Neanderthals. *Nature, 388*, 225–226.

Weisdorf, J. L. (2005). From foraging to farming: Explaining the Neolithic revolution. *Journal of Economic Surveys, 19*, 561–586.

Wilkins, W. K., & Wakefield, J.(1995). Brain evolution and neurolinguistic preconditions. *Behavioral and Brain Sciences, 18*, 161–226.

Williamson, S., & Wright, R. (1993). Barter and monetary exchange under private information. *American Economic Review, 84*,104–123.