Inter-generational transitions in socio-technical systems: The case of mobile communications

Shahzad Ansari\textsuperscript{a,}\textsuperscript{*}, Raghu Garud\textsuperscript{b,1}

\textsuperscript{a} Department of Strategic Management and Business Environment, RSM Erasmus University, Burg. Oudlaan 50, 3062 PA Rotterdam, The Netherlands
\textsuperscript{b} The Pennsylvania State University, 430 Business Building, University Park, PA 16802-3603, United States

\textbf{A R T I C L E  I N F O}

Article history:
Received 15 November 2007
Received in revised form 5 October 2008
Accepted 15 November 2008
Available online 30 December 2008

Keywords:
Technological transitions
Socio-technical system
Collateral technologies
Mobile communications
S-curve

\textbf{A B S T R A C T}

Many technology studies have conceptualized transitions between technological generations as a series of S-curve performance improvements over time. Surprisingly, the interregnum between successive generations has received little attention. To understand what happens in the interregnum, we study the transition from 2G to 3G in mobile communications. Our study identifies the presence of forces for both change and continuity across heterogeneous social and technical elements shaping an uneven transition between 2G and 3G mobile communications technology platforms. Unanticipated misalignments and asynchronies that emerged during the journey shifted the incentives of the various actors involved to participate. Based on these observations, we offer several conjectures as to the dynamics that can give rise to temporal discords during inter-generational technological transitions.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Many technology studies are predicated on Schumpeter’s (1942) perspective wherein waves of discontinuous technological change unfold within an overall process of creative destruction (cf. Christensen, 1997; Hill and Rothaermel, 2003; Tushman and Anderson, 1986; Utterback, 1994). With an implicit assumption about the inevitability of technological trajectories, technological evolution has been conceptualized as proceeding in a sequential and progressive manner prescribing an S-shaped curve as seen, for instance, in performance improvements in desktop memory (Christensen, 1997). The S-curve hypothesis suggests that the performance of a technology, slow at first, accelerates over time, finally flattening out to be supplanted by a new technology with its own S-curve.\textsuperscript{2}

Despite the seductive appeal of this framing and the clear implications that actors should shift investments from a mature to a new technology at an “inflection point” (point of saturation) on the S-curve (e.g. Christensen, 1997; Grove, 1996), those who have examined the micro-processes of technology transitions suggest that there is more than a simple shift from one technological trajectory to another (Geels, 2002, 2005; Sood and Tellis, 2005). For instance, old technologies can prove to be surprisingly resilient and may not simply get eclipsed by new ones (Henderson, 1995). In addition, the effective functioning of any technology requires the creation of co-specialized assets (Teece, 1986), a facet of technology transition that may not unfold as anticipated. Complicating matters, different user groups may have different interpretations of a technology-in-use, thereby impacting any transition (Pinch and Bijker, 1984). Moreover, with transitions that involve platforms with multi-sided markets\textsuperscript{3} (Boudreau, 2006; Evans et al., 2006; Rochet and Tirole, 2003), coordinating and balancing the diverse strategic and economic interests of multiple sets of users further increases the complexity of the transition process.

These observations suggest that it would be productive to examine the processes that unfold during a transition. To do so, we adopt a socio-technical perspective (Bijker et al., 1987) to examine the transition from the second (2G) to the third (3G) generation in the field of mobile communications. By examining the interactions across heterogeneous social and technical elements over time, we

\textsuperscript{*} Corresponding author. Tel.: +31 10 408 1996; fax: +31 10 408 9013.
E-mail addresses: sansari@rsm.nl (S. Ansari), rgarud@psu.edu (R. Garud).

\textsuperscript{1} Tel.: +1 814 863 4534; fax: +1 814 863 7261.

\textsuperscript{2} Such improvements correspond with the diffusion rate of novel technologies and innovations over time (Rogers, 2003). Only a few actors adopt an innovation at first. Then, the adoption rate increases sharply with large numbers adopting followed by a slowing down as laggards finally adopt the innovation.

\textsuperscript{3} We are grateful to an anonymous Research Policy reviewer for drawing our attention to this emerging literature. An example of a multi-sided platform would be a videogame platform, such as Nintendo that not only needs to attract users in order for game developers to design games for its platform but also needs games from game developers to induce users to buy and use Nintendo’s console.
find that transitions are complex and non-linear processes wherein unanticipated misalignments and asynchronies that emerge during the journey can shift the incentives of the various actors involved to participate. For instance, the viability of investments may get problematized during implementation resulting in deterring complementors and co-specialized assets providers, user preferences may get clarified in-use as they interpret technologies in unanticipated ways and collateral technologies may enter the picture to reduce the performance gap between the old and the new. All these developments can conspire to change parameters even as a transition unfolds. As a result, technological transitions may not follow “natural trajectories or “a set logic in advance” (Callon, 2008).

2. Theoretical framework

Schumpeter’s (1934, 1942) seminal work on creative destruction has inspired a rich body of literature on technological cycles, discontinuities, paradigms and trajectories (Dosi, 1982; Tushman and Anderson, 1986; Utterback, 1994). A popular hypothesis on technological evolution as captured by the image of an S-curve continues to attract much scholarly attention (Christensen, 1997; Sood and Tellis, 2004). Such imagery suggests a sort of inevitability about technological trajectories where, despite fierce battles, new technologies eventually eclipse older ones.

However, while an S-curve may be useful in depicting the outcomes of a battle between technological generations once it has been fought, it is less useful in understanding issues that arise during transitions (Latour, 1991). Empirical research on technological transitions has problematized the S-curve by showing that old technologies may persist and even improve following long periods of dormancy (Henderson, 1995; Lawless and Anderson, 1996; Sood and Tellis, 2005). That technological transitions may not be linear has also been noted by those who have taken a systemic view on technological platforms (Gawer and Henderson, 2007). With such a view, it becomes clear that different inter-related co-specialized assets (Teece, 1986) of any platform may emerge at different rates, thereby generating critical “bottlenecks” along the way (Rosenberg, 1982).

Social construction of technological systems scholars (e.g. Bijker et al., 1987) have further opened up the black box of technological transitions by examining not only technical but also social elements at play during transitions. For instance, associated with any technological platform are actors from the supply, demand and institutional domains (Garud and Karnoe, 2003; Geels, 2005). Supply-side actors include a constellation of organizations that are involved in the design, production and distribution of goods and services. On the demand side, there may be multiple users, a facet of complex technological platforms that recent research on multi-sided markets highlights (Boudreau, 2005, 2006; Evans et al., 2006; Rochet and Tirole, 2003). Mediating many of the interactions between and across these groups are institutional actors such as regulators, media critics, analysts and the like (Kemp et al., 2001; Lounsbury and Glynn, 2001; Pollock and Rindova, 2003; Zuckerman, 1999).

An expanded view of technological platforms highlights why technological transitions may not be straightforward. For instance, from the supply side, even simple changes can hold important implications for the re-distribution of competencies across production networks (Garud and Munir, 2008; Glasmeyer, 1991). Consequently, changes might be resisted. This is likely to be true from the demand side as well, especially with respect to transitions that either disturb or attempt to create a new equilibrium between users on multiple sides of a platform (Boudreau, 2006; Rochet and Tirole, 2003). Even with single markets, linear transition can easily be upset because of the different and at times unanticipated ways in which users may interpret a technology (cf. Pinch and Bijker, 1984).

Besides supply-side and demand-side forces, institutional forces such as existing methods of evaluation, the legitimacy of incumbent technologies and prevailing rules and regulations may encourage the status quo rather than change (Hargadon and Douglas, 2001; Kemp et al., 2001).

How might those who want to bring about change counter these forces for continuity? One way is to create a compelling vision of the future (Brown et al., 2000). Especially during early stages of technology development, collective beliefs and expectations about the future can play a crucial part in defining roles, building interests and constructing mutually binding obligations among a set of interdependent constituents operating in different domains (Van Lente, 2000). Through claims and promises, expectations create a rhetorical space that over time evolves into a reality of requirements, mutual obligations and “emerging irreversibilities” (Callon, 1991). By structuring the activities of a distributed network of actors within a socio-technical system, expectations of the future can come to play a self-fulfilling role just as financial instruments can play a “performative” role in constituting markets (Callon, 1998; MacKenzie and Millo, 2003).

How forces for continuity and change, both present during any attempted transition, play out can have a profound impact on the specific path along which a transition unfolds. It is this facet of change that we are interested in examining in this paper. For instance, what are the temporal discords that emerge during a transition? How do they dynamically change the relationships among members in a socio-technical system with diverse interests and perspectives? We address these questions by providing an account of the processes involved in the transition between the second generation (2G) and third generation (3G) of the mobile communications platform. But first, we offer a brief overview of our research site and the methods that we have employed to conduct this study.

3. Research site and methods

A study of mobile communications from 1999 to 2006 represents a particularly revelatory case for examining the dynamics of inter-generational transition in the wake of a technological discontinuity – the advanced third generation (3G) technology. From a performance standpoint, the field of mobile communications has seen three main generations from the early 1980s to the present, from voice-centric 1G and 2G platforms to voice and data-centric 3G platforms along with an intermediate 2.5G platform and discussions about a futuristic 4G. Fig. 1 depicts how the performance has improved over time.

Two factors shaped our decision to examine this field. First, it is characterized by systemic properties (Katz and Shapiro, 1985) with interdependent constituents such as network operators, equipment manufacturer, content providers, regulators, and users, each guided by different frames, interests and motivations, vis-à-vis the transition. Second, events relating to technological transitions in this domain are well documented. This makes it possible to trace the developments of multiple inter-related constituents over time and to theorize about the transition.

3.1. Data sources

Data concerning the actions of various constituents were acquired from three sources. First, we interviewed several key exec-

---

4 The overall movement is a “promise-agenda-requirement” conversion process (Van Lente, 2000:61). An expectation statement indicates promising aspects of technological developments, allocate roles, structure activities, invokes responses and creates an agenda for activities. When such an agenda becomes accepted, it may then lead to a more detailed specification of technological requirements and resources.
utives in many of the firms involved in adopting 3G technologies or those choosing not to adopt 3G. Similarly, we interviewed several intermediaries regarding the roles they had played in this migratory process. Second, we participated in several conferences such as the 3GSM conference and the Financial Times World Mobile Communications Conference that were well attended by key stakeholders in the field. Third, we examined relevant articles from 1997 to 2007 from several sources (e.g. Financial Times, The Economist, Wireless Review, Wireless News). We accessed data from company websites, internal reports, and analysts’ reports on mobile communications (e.g. Credit Suisse First Boston, Lehman Brothers, Merrill Lynch & Co and HSBC), dedicated websites (e.g. www.telecom.com, 3Gnews.co.uk, Zdnetuk.co.uk), online data-bases (e.g. Lexis-Nexis, ProQuest, Reuters) and academic journal articles. We also obtained data from regulatory agencies such as the UK telecommunication regulatory agency Ofcom, the Universal Mobile Telecommunications Services (UMTS) forum (an organization representing the industry, regulators and other government institutions), OMA (Open Mobile Alliance) (consisting of over 300 companies representing key constituents) and telecom consultancy companies. Finally, we made use of several books on mobile communications.

3.2. Data analysis

We first coded the data by assigning labels to text units (sentences or paragraph) (Miles and Huberman, 1994; Strauss and Corbin, 1990). After we identified potential categories, we grouped initial text units into a smaller number of aggregated themes in order to understand the underlying patterns. By systematically analyzing all interview transcripts and archival documents, we identified a series of patterns vis-à-vis the forces of change and continuity.

We present our data as a narrative to understand how contradictory forces of change and continuity mediate technological transitions. Narratives provide a temporal ordering of events involving both actors and technologies (Pentland, 1999). Surface-level details contextualize observations while underlying driving forces provide a way to apply learning from one setting to another setting (Bruner, 1990). Following this approach, we recount the surface-level details based on an analysis of salient events that unfolded (Miles and Huberman, 1994). Then, we examine the underlying driving forces paying specific attention to forces for change and continuity. For instance, in the mobile communications case, we found that early forces for change were dampened by delays in the development of co-specialized assets and that the emergence of collateral technologies changed the value proposition for constituents even as the transition was unfolding.

4. Technological evolution in mobile communications

On April 19, 2000, the UK government held the first 3G spectrum auctions in Europe, raising $35 billion from the sale of five licenses. Soon thereafter, the spectrum auction in Germany raised almost $50 billion. While subsequent auction proceeds from spectrum auctions in Europe were more modest, the combined revenues from the sale of licenses for 3G mobile technologies raised more than $100 billion. Not all countries chose the auction mechanism to allocate spectrum licenses. For instance, Finland and Sweden allocated these licenses through an administrative mechanism described as a ‘beauty contest’ where spectrum licenses were simply awarded to ‘deserving’ candidates (mobile operators) but who nevertheless were required to make additional investments in developing 3G networks. Yet the enthusiasm for 3G was no less even in these countries. The governments’ decision to sell or award valuable spectrum had sparked the creation of a new market for the next generation mobile communications.

Five years down the road, the use of products and services based on 3G technologies remained well below expectations. Many mobile operators had to write down the value of their spectrum licenses while some even returned them to the state at considerable loss. Others delayed rolling out 3G networks. This was true for even those operators who had not incurred expenses to obtain licenses because the licenses had been awarded by their governments. Despite the optimism that the third generation mobile technologies had generated at the time of spectrum allocations, by 2005, of the almost 2 billion mobile customers worldwide, just about 2% had migrated to the 3G platform – far below initial estimates. Since then, operators (that offer 3G) have made 3G functionalities available to almost a quarter of their subscriber base (this includes subscribers who have simply signed up with a 3G service provider), but users have shown little interest in using specific ‘high bandwidth’ data-based services (e.g. video telephony) that 3G offers (Seybold, 2004; www.telecom.com). See Table 1 for subscriber numbers for various technological generations.

In the interim, many operators decided to focus on upgrading the previous generation 2G technologies to what came to be known as 2.5G, a hybrid that had elements of both 2G and 3G. These operators included those who had obtained a 3G license and viewed 2.5G as a “stepping stone” to 3G (Brodsky, 1998) as well as those who did not get a 3G license and viewed 2.5G as an “end in itself” (Pesola, 2001). This interim ‘solution’ offered almost all of the benefits that 3G had promised at a fraction of the cost and without the need for new spectrum licenses and mobile infrastructure. In a sense, users’ emerging preferences appeared to congeal around the functionality 2.5G could offer vis-à-vis data-based content (Pesola, 2001). As the hybrid label 2.5G suggests, instead of a smooth transition from

![Fig. 1. Neither fish nor fowl but 2.5G. Instead of a smooth progression from the second generation (2G) to the third generation (3G), an interim two and a half generation (2.5G) – a hybrid containing elements of both 2G and 3G and offering a performance level above 2G but below 3G – emerged that was able to close the performance gap between 2G and 3G and impact the transition from 2G to 3G.](image-url)
the second to the third generation, the system settled down somewhere in-between, incorporating facets from both generations. To understand how and why this happened, we need to look beyond the technologies at play in order to examine the forces that emerged during the transition from different parts of the socio-technical system for mobile communications.

4.1. The mobile socio-technical system

At the time of spectrum allocations, the mobile system was locked into a voice-centric 2G platform – essentially the wireless equivalent of fixed-line telephony – that had reached saturation in terms of penetration levels. 3G could change the rules of engagement by making mobile communications voice and data-centric. With phones becoming ‘hybrid’ devices (offering voice, video, mobile TV, Internet access and so on), there were changes unfolding in the “industry architecture” – template of agents and assets that circumscribe the contours of an industry (Jacobides et al., 2006) and boundaries of the mobile platform were becoming increasingly “plastic” (Boudreau, 2005) as mobile communications entered the realm of consumer electronics and entertainment. This required not only new core technologies but also several co-specialized assets (Teece, 1986). Among these assets were new generation mobile handsets from manufacturers (that needed to be compatible with previous generation mobile phones), new base stations and masts from infrastructure providers for the transmission of 3G signals and compelling applications or mobile ‘content’ (video games, websites, etc.) from application developers for the 3G platform.

The 2G platform only offered voice-based services with all end-user revenues being appropriated by the mobile operators. The new market around the 3G platform was premised around two kinds of users – content providers (such as banks and entertainment companies to develop new kinds of ‘mobile’ services) and end-users who would benefit from the new functionalities being offered. 3G effectively stood to make mobile communications a “multi-sided market” where mobile operators would need to serve multiple complementary sets of users (Rochet and Tirole, 2003). Since spectrum is a state-controlled resource, the field also involved regulators. Similarly, intermediaries – analysts and media – played an important role in shaping perceived realities. Besides, environmental groups concerned about the potential radiation hazards from the 3G platform were also important in influencing the transition process. Before we elaborate further on the roles of various actors in the mobile communications field, we first provide an overview of the essential characteristics of 3G technologies and the advantages it promised over the preceding 2G technologies.

4.2. Key characteristics of the third generation (3G) wireless technology

In June 1999, nine global leaders in the mobile telephone formed an industry group called 3GPP to develop the technology required for the next generation (3G) of wireless communications (Taylor, 1999). As against the globally dominant GSM (Global System Mobile standard) in 2G technologies, 3G technologies used the CDMA (Code Division Multiple Access) radio interface standard. CDMA is far more efficient in ‘farming’ available spectrum by assigning a unique code for each frequency channel, thereby providing more traffic per megahertz of spectrum (Funk and Methe, 2001). CDMA has two versions; the European sponsored Wideband CDMA (UMTS) and the US centered CDMA2000. Most 2G technologies, including the European-led GSM standard, was based on TDMA (Time Division Multiple Access) – a technology that creates multiple access channels for subscribers by spacing frequencies in time. 3G technologies not only made more efficient use of a wider band of spectrum, but also had the capability to transmit more data at a much faster speed than 2G.

Besides these differences in core radio interface technologies, 3G also offered a more efficient way of transmitting wireless data. All 3G technologies use ‘packet based’ systems for sending data – a technology also employed for data transmissions over Internet (Martellini, 2008). The file to be transmitted is split up into smaller units or packets that are then reassembled at the file’s destination. Spectrum usage is more efficient because the channel remains available to other users during the connection between two users. In contrast, circuit-switched technology that was employed in voice-centric 2G technologies sets up a dedicated connection between two callers for the entire duration of the communication, thereby creating inefficiencies in spectrum utilization. While the core 3G technology, CDMA, and packet switching allow for better utilization of spectrum, handling large amounts of data also required a higher spectrum capacity or bandwidth. Spectrum being a scarce and valuable resource (also used for example by national defense services) meant that several operators had to obtain licenses to access 3G spectrum. To understand the enthusiasm for 3G technology, we need to look at the antecedents to 3G spectrum allocations. This includes the context in which the allocations took place and the activities of various constituents in the run up to these allocations.

4.3. The enabling context

As with other technological fields, actors in the field of mobile communications responded to the broader context shaping change. Mobile communications took its cues from ‘broadband’ where fixed data services were growing at more than 200% a year (Deloitte Report, 1999). Taking note, the Director of data strategy at Lucent Technologies stated: “The same market pull will hit the wireless industry and with third generation standards, mobile will be ready to meet demand when it arises” (quoted in Awdle, 1999). CEOs of major operators, such as Vodafone believed that “just as voice is moving to wireless, so will data and the Internet, for the same reason: convenience” (quoted in The Economist, 1999).

Projections that growth in data-based services would soon impact the wireless domain, led many in the industry to believe that the voice-only functionality offered by 2G would not be enough. It was in this context that 3G promising the functionality required to seamlessly surf the web on mobile phones was proposed. Since almost everyone owned a mobile phone, 3G would make the Internet mainstream and widely accessible by eliminating the need for users to own expensive laptops and PCs (The Economist, 2004a). Furthermore, 3G was seen as a platform that could spur demand for online shopping (e-commerce) by enabling online purchases through the mobile phone (m-commerce) (The Economist, 2001). In essence, 3G promised to make mobile communications a multi-sided market by expanding the boundaries of mobile telephony to include information and entertainment services. Switching to 3G was thus seen as the most viable growth strategy for mobile operators who widely agreed that the old 2G technology had reached its “natural performance limit” (Foster, 1986; Sahal, 1985).

5 3G networks operate at higher frequencies and the range of radio waves decreases as frequency increases. 3G networks require more cells than 2G networks (Newing, 2001). This required the construction of additional masts and base stations in cities as well as more powerful handsets than those used in 2G technologies. Many believe that these components posed significant radiation hazards.

6 One exception was the use of CDMA in incumbent 2G systems, such as the CDMA-One (IS-95) that used 2G spectrum. However, with the TDMA-based GSM being the de facto global standard for 2G, the use of CDMA-One remained largely limited to North America and parts of Asia (Funk and Methe, 2001).
4.4. Firms’ reasons for switching to 3G

The end of the voice-based mobile boom7 (Sammut-Bonnici, 2005; The Economist, 1999) meant that revenues from traditional voice-based telephony had reached almost saturation levels and that there was little scope in adding more users to the network.

"International and long-distance call prices are likely to continue to fall rapidly over the next two to five years . . . this is a bleak outlook for any operator requiring a five-to-10-year return on capital employed. The mobile sector seems just fraught with danger" (Analysys Report, 1999).

The only way to grow then was to increase the usage of the network or the ARPU (average revenue per user). 3G was expected to create a huge surge in revenues from data-based services that would require much greater bandwidth than what existing, overburdened and low bandwidth 2G technologies were capable of delivering. Eager to tap new sources of revenues, it seemed logical for mobile operators to switch to 3G and obtain licenses for accessing additional spectrum.

Besides the perceived expectation of an imminent decline in traditional 2G services, there were several other reasons for firms’ decision to migrate to 3G. First, as discussed above, given that financial markets were expected to reward firms that made this transition, eager shareholders of mobile firms exerted considerable pressure on management to acquire 3G licenses. Second, as 3G promised a single, unified global standard,8 there was a temptation to create a viable global ‘footprint’ with seamless connectivity. This required a license at multiple locations as firms hoped to build sufficient size and scale and achieve higher bargaining power with suppliers to drive not just pan-European but global branding. Third, with 3G, mobile operators could “monetize the Internet” and earn revenues through web-based services9 (The Economist, 2001).

Mobile operators believed that they had an important advantage over Internet service providers and fixed-line companies as noted by an interviewee just before the UK license auction:

“I think the mobile operator with the capability to bill for telephony and also for ‘service packages’ has a big advantage over Internet service providers with no direct billing relationships and even the fixed operators who normally bill a home or an enterprise rather than an employee or a person” (MD of Hutchison, UK).

While there seemed many compelling reasons for firms to invest in 3G, several other actors played an important role in promoting 3G. We discuss the role of these constituents in fuelling optimism for the proposed new technology.

4.5. Support from other constituents

Equipment manufacturers were eager that mobile operators make the switch to 3G given its perceived potential. With their 2G product lines nearing obsolescence, manufacturers, such as Nokia and Ericsson stood to considerably benefit from the development of the 3G market. Indeed, these companies were among the first to come up with the term ‘3G’ to describe the technologies as a new generation.

Such forward-looking thinking was not restricted only to business but also prevailed in national economies. Our analysis suggests that European politicians eager to adapt themselves to the conditions of the new economy had strong motivations to raise optimism for 3G – seen as a European-led global standard that would maintain Europe’s lead over US in mobile communications. 3G was also expected to benefit the economy by generating new business ventures and jobs and unleashing a wave of innovations through the convergence of media, telephony and the Internet. Indeed, our review of the reports on the governments’ decision to designate 3G spectrum suggested a strong preference for the proposed new technology. Despite the “large uncertainty” about demand for 3G services, the price elasticity and the nature of competition in future 3G markets, these reports painted an optimistic scenario for 3G as a technology (Borgers and Dustmann, 2003:228). In the UK, for instance, reports from Ovum Ltd and Quotient Communications Ltd (1998) based their analysis on the figures for London – UK’s largest city with a dense network and heaviest usage of mobile services. London was chosen to ensure that the demand for 3G services would not be underestimated. Many of these reports concluded that early users of 3G services would be relatively ‘insensitive’ to cost and that demand would be independent of price considerations (Communicator, 1999:6). The government then engaged in a sustained marketing campaign to attract entrants (Binmore and Klemperer, 2002), describing the technology as truly “revolutionary.” A UK official’s quote is indicative of the state’s support for 3G.

“Third Generation (3G) mobile technology will transform how we access information, creating applications and opportunities, we are only beginning to imagine. Along with Interactive TV, it will bring the Internet to people who would never dream of having a computer at home” (Patricia Hewitt, Small Business and E-Commerce Minister, quoted in Financial Times, 1999).

Besides regulators, infomediaries such as investment bankers, and analysts also upped the ante for 3G and linked the ratings of mobile companies to their engagement in 3G projects. Having severely underestimated10 the tremendous growth during the second generation, many analysts were keen to atone for the mistake. For instance, despite earlier revenue forecasts unable to justify the increasing license price, an advisor to a mobile operator under pressure to come up with up with more optimistic findings, subsequently revised its estimates (Lehman Brothers Report, 1998). The press also contributed to the positive framing of 3G, as indicated by the quote:

“3G will be on a different plane. For starters, the data rate is phenomenal – 2 M bytes per second if you are in a ‘cell’ close to a transmitter, or 384k bytes for normal mobile use. At that rate, videophones become an everyday reality, along with television and radio on demand and multimedia Internet content surpassing anything available to the average consumer today” (The Sunday Times, 1999).

Positive anticipations for the proposed new technology at the time appeared to crowd out the few voices of dissent regarding 3G’s prospects. A wireless expert noted that, despite the enth-

---

7 The creation of a European-led dominant global mobile standard has resulted in phenomenal growth rates in mobile telephony, with an average rate of over 60% from 1995 to 2000.
8 The regulators wanted to avoid the patterns of global fragmentation that had characterized the market during the first and second generations. North America, Japan, and Europe retreated into separate camps resulting in an isolated, balkanized landscape of incompatible technologies, especially in the US.
9 Internet service providers face major challenges in making money from web-based services (news, weather reports, etc.) that are mostly ‘free,’ given that most do not enjoy direct billing relationships with end customers. Mobile operators, in contrast, were in a unique position to charge for such information.
10 For instance, in a market research report for AT&T during the first generation mobile communication technology during the mid-1980s, the firm McKinsey had come to the conclusion that the cell phone market would not be a profitable option as the worldwide market potential would be only around 900,000 phones by the year 2000 (The Economist, 1999). However, the number of cellular phones in that year exceeded 400 million globally.
siasm, there were “applications yet to be identified, technologies yet to be proved and markets yet to be served. Until those tasks are accomplished, the 3G wireless standards competition is just another ongoing debate” (Brodsky, 1998). Many analysts noted that there wasn’t enough market data to prove that the demand was really out there for wireless multi-media (Jensen, 1999) and that the mobile operators could lose a “whole lot of money” (Sobti et al., 1998). We quote from an analyst’s report from 1998.

“All of that assumes we know what we are going to do with 3G and we don’t. Nobody really has an application or service to offer on 3G that can’t be offered on 2G today. They will talk to you about video this and video that and high-speed imaging, but there isn’t anything yet” (Hersche Shoestock Associates Report, 1998).

4.6. The 3G spectrum allocations

Given the collective framing around 3G, it was no wonder that many operators were very eager to obtain licenses to access a specific bandwidth on the spectrum that governments designated and put up for allocation in 2000–2001. In some cases, these licenses were awarded administratively. Others ended up paying huge prices – a total of over 100 billion Euros. In both cases there were stringent rollout obligations attached to spectrum licenses. For instance, the UK regulators required each licensee to cover at least 80% of the UK population with 3G services by the end of 2007 (Preliminary Information Memorandum, 1999). As 3G required a completely new network, meeting this obligation in sparse demographic distributions and remote terrains would require huge further investments. The total projected investment for rolling out 3G network infrastructure from 2001 to 2010 in Europe was about another 250 billion Euros. Given the promise of new markets in mobile communications, mobile firms nevertheless made huge investments in their commitment to 3G.11

4.7. 3G’s stumble

More than five years down the road, the migration had not occurred as intended despite a concerted ‘technological push’ from many operators. Analysts’ forecast at the time of license allocations had been that the technology would be up and running by 2003–2004 and that it would generate significant rents by 2005 (Lehman Brothers Report, 1998). Even at the end of 2005, 3G penetration levels and use remained well below expectations (Sylvers, 2007). We explain some of the factors that shaped this outcome.

4.7.1. Resource constraints

Firms found themselves under pressure from bankers and shareholders to reduce the debts that they had accumulated. With the increasing reluctance of financial institutions to lend more to an already embattled sector, funds to build new 3G infrastructure were in short supply. To make matters worse, the investment and operating costs for a 3G network turned out to be much higher than initial projections (Burgelman and Carat, 2003) putting further strain on firm resources. Not surprisingly, several firms, such as the German operators Mobilcom and Quam decided to simply return the license to the regulator while others had to scale back their targets, withdraw from unattractive markets and delay or even abandon 3G rollouts in violation of their regulatory commitments. To make matters worse, the demand for 3G services turned out to be well below expectations as end users made little use of the added 3G functionalities that were being offered.

4.7.2. The role of users

Users did not enact the 3G script as encoded by its proponents. A Nokia executive noted: “The industry has not been looking at the user. It’s been looking at its own navel” (Baker and Clifford, 2002). Most research indicated that people were still limiting their use of the mobile for ‘phatic’ or social communication as against consuming third party content (such as buying goods online) possible with 3G (Wilson, 2006). Indeed, social communication (talking and text messaging) still accounts for well over 90% of operator revenues. In a recent survey conducted among mobile users in Europe, 44% of respondents said that they would not use their mobile phones other than to make ordinary calls (Wall Street Journal, 2004). Thus, while 3G was conceptualized as a tool to extend the mobile platform beyond telephony and offer instantaneous information and entertainment services, the essential use of the platform continued to be social networking with very limited uptake of digital content – a dynamic that some derisively noted as indicating mobile “discontent” (Wilson, 2006). The persistent low demand for 3G services prevented positive consumption externalities from kicking in (Economides, 1996). This deterred potential users whose utility from adopting 3G services depended on the number of other 3G subscribers.

Furthermore, since new 3G services were dependent on complementary assets such as compatible handsets, delays and glitches further retarded 3G’s adoption. The UK head of Hutchison 3G, a mobile operator, complained of a lack of handsets.

“Heavy handsets with low battery life did not help our cause... Nokia (the leading manufacturer) adopted a wait and see policy for investing in 3G handsets and delayed deliveries that severely disrupted our operations” (MD of Hutchison 3G, UK).

As 3G needed complementors developing content for the network (e.g. mobile banking) banking software had to be customized for the rapidly evolving 3G mobile handsets. Developing such software required changes every time a new handset was released, dampening the availability of such content. The lack of compelling content further slowed down 3G’s diffusion. Mobile operators, given their financial situation and not used to sharing revenues in the 2G world, wanted to control the largest possible part of the value chain by limiting third party access to their platforms in the hope of increasing revenues. This ‘closed’ or ‘walled garden’ model rather than an ‘open’ model for access to platforms that operators embraced deterred not only complementors but also end-users who were accustomed to unrestricted content on the Internet (du Preez and Pistorius, 2003). This negatively affected the development of content for the mobile platform (Wallage, 2005). Furthermore, in Europe’s multicultural and multilingual environment, several content providers (e.g. newspapers) had to target smaller local audiences, a market that did not offer scale advantages. In some cases, content providers simply could not afford to deliver mobile content for such “small” target audiences (Andersen, 1999). Given low volumes, content providers could not adopt the advertising-based business model of ‘free content’ as was often the case with the Internet. Instead, in the case of mobile communications they wanted their content to be profitable on a standalone basis.

To complicate matters, financially constrained firms with low additional revenues from 3G services were unable to adequately invest in marketing 3G services to gain end-user acceptance that

11 It is worth noting regional differences in 3G spectrum allocations. For instance, U.S. wireless operators were in no hurry to move ahead with the implementation of 3G in sharp contrast to the enthusiasm shown by several of their European and Asian counterparts. An executive from a major operator noted: “I can’t get too excited until I know what the capital cost is and what my revenue stream will be. Really, I just don’t see the need to go headlong into 3G at this point” (quoted in Wireless Data News, 1999). In fact, at the time, FCC and other U.S. government agencies had yet to lay the groundwork for assigning spectrum that would be used for 3G services.
further retarded the diffusion of the new technology. Another important factor in the slow uptake were ongoing battles between the two competing 3G radio interface standards for the 3G platform – the US led CDMA2000 and the European led WCDMA – even though 3G was meant to be a unified global standard. Furthermore, infomediaries – analysts and media – who had earlier promoted the new technology now turned decisively against 3G. Moreover, 3G faced resistance from environmental groups concerned about suspected harmful effects of 3G systems and was also threatened by the development of substitute technologies.

4.7.3. The role of other constituents

Many analysts who had promoted the new technology at the time of spectrum allocations started to speak of much longer (up to 30 years) payback periods for 3G investments and now opined that firms had exposed themselves to a high risk of premature technology introduction (McClelland, 2003). An analyst argued that mobile firms had not taken into account the fact that the demand for mobile access to ‘content’ through 3G was “derived” from the demand for services delivered over the Internet (Ure, 2000). This was unlike the previous voice-focused networks (1G and 2G) in which demand for services delivered over the Internet was “intrinsic” to the mobile networks.

Not surprisingly, the media that had vociferously supported 3G, just prior to spectrum allocations did a volte-face, subsequently labeling 3G “Universal Mobile Telecommunication Services” (UMTS) as “Unproven Market, Technology and Services.” Even some national governments echoed the pessimistic scenario with the Finnish minister of transport and communications calling 3G auctions “the biggest industrial political failure since the Second World War” (Klemperer, 2002). Environmental groups also put up strong resistance to the building of the much needed new base stations and 3G masts (transmitters), in view of their perceived links with radiation hazards. In short, several constituents contributed to building collective ‘negativity’ around 3G that heightened uncertainties regarding the new technology and retarded its diffusion.

To make matters worse, a number of alternative technologies began casting a long shadow over the prospects of 3G. One of these alternatives was Wireless Fidelity (WiFi) that could siphon data traffic away from 3G networks by providing access to mobile information in designated locations known as ‘hotspots,’ such as airports, coffee shops, or universities. There was speculation in the industry that all the hotspots would be linked in the future to create a seamless universe with “blanket coverage” for all users thereby making 3G redundant (The Economist, 2004b). It was even feared that, rather than generating new sources of revenue as was originally expected, 3G technology could end up merely as a way for mobile operators to boost their capacity for traditional voice calls in the overloaded parts of their 2G networks (The Economist, 2001). In short, 3G was no longer regarded as a paradigm shift in mobile communications that would open up a goldmine of new revenues for mobile operators. As 3G continued to disappoint, many firms in the industry began to explore the possibilities for stretching the capacity of 2G networks, a possibility that had hardly received any attention during the early days of 3G when a generational shift was being considered.

4.8. Extension of the old technology

In the meantime, upgrades to 2G technologies, labeled as 2.5G or GPRS (General Packet Radio Services), were being achieved by ‘bolting on’ packet-switching data transmission technology onto existing 2G technologies. Although packet-switching technology was available even at the time of 2G (GSM), it was given “fairly marginal billing” at the time and was not incorporated into the GSM system in the mid 1990s (Pesola, 2001). According to Mannings and Cosier (2001), this is because firms in the mid 1990s had a voice-centric “telco mindset” and arguably didn’t anticipate the potential fallout on mobile telephony from developments in the Internet. 2.5G essentially allowed faster connections to the Internet via the mobile phone while still operating on 2G spectrum. 2.5G could handle data transmission speeds between 33.6 and 128 kbps, as against a maximum of 2000 kbps for 3G, thereby lifting the performance to a level which is in the middle range of 2G and 3G. As an example, downloading a song with a 2.5G network can take over a minute as against 5 s on a 3G network. 2.5G was like introducing a double-decker bus on an existing one-lane highway (voice plus data) rather than build a new 4-lane highway (3G) (Newing, 2001). Also, 2.5G services came with reduced price tags compared to 3G as upgrading existing networks entailed no additional costs for obtaining 3G licenses and for building new 3G networks.

2.5G was being pursued by two types of mobile firms – firms that did not obtain a 3G license and firms that obtained 3G licenses and also invested in 3G rollouts. For the former, this strategy was an “end in itself” (Pesola, 2001). It represented a way to remain competitive in the mobile industry and hurt 3G’s prospects by showing the viability of improving existing networks. Firms without 3G licenses, such as the Danish operators, Sonofon and Telias focused on 2.5G to offer many services that were supposed to be ‘3G’ services and many executives suggested that consumers have no need for 3G.

In contrast, for firms with 3G licenses who had to delay their 3G launches beyond initial promises (such as Vodafone), ‘backpedalling’ to 2.5G represented a “stepping stone” to creating demand for ‘full blown’ 3G (Sammut-Bonnci, 2005). However, firms confronted a dilemma – preventing the cannibalization of their customer base in a near-saturated market while, at the same time, also requiring them to sign up for new 3G contracts aimed at increasing the ARPU.

As the system temporarily settled down around 2.5G, many operators now acknowledged that the transition to 3G was unlikely to be momentous and began to regard 3G as evolutionary rather than revolutionary. Indeed, as the CEO of BT Cellnet (later known as mmO2) stated in 2003: “In terms of the kinds of applications our customers are going to use, I don’t think there is a huge difference between 2.5G and 3G” (Telecom Review – Conference proceedings in the 3GSM World Congress in 2003). The situation actually was far more complex than having to decide between two rival technologies. According to the VP of Semico Research Corp. “It is going to be a case of supporting 2G, 2.5G and 3G, all at the same time. This isn’t going to be a nice, smooth transition.”

Worse was the prospect that firms that chose to upgrade to 2.5G could bypass 3G and make a direct leap for 4G (The Economist, 2003). These 4G networks (also known as Beyond 3G), based on new technologies such as improved modulation and relaying were expected to deliver voice, data and full motion ‘streaming’ video on mobile devices at over ten times the speed of 3G. A respondent noted that given what 4G could bring, 3G could well be a “stillborn” technology and the third generation a “lost generation” (former head of technology of UK telecom NTL).

5. Towards a dynamic perspective on technological transitions

We have built on the work of scholars arguing for the need to take an expanded socio-technical perspective in conceptualizing technological transitions (Bijker et al., 1987). By disaggregating various elements of the mobile communications socio-technical system, we identified several forces that shaped the transition between 2G and 3G mobile communications technologies. Forces for change were mediated by dampening forces stemming from delays in implementation, developments in collateral technologies and concealing of user preferences around a set of functionalities below what had been projected for the new technology. These forces conspired to
change preferences and incentives of the constituents involved even as the transition was unfolding to disrupt the carefully constructed connections that had been framed to spark the ‘intended’ transition path. We discuss how these dynamics played out in greater detail so as to offer some conjectures about transition processes.

5.1. Forces for change

Since additional spectrum was a prerequisite for migrating to the next generation technology, the government’s decision to allocate spectrum was performative in generating forces for change by creating a multi-sided market (Boudreau, 2006; Evans et al., 2006; Rochet and Tirole, 2003) in mobile communications. By designating a specific time frame and by allocating a particular bandwidth for 3G spectrum, national regulators, in effect, ‘imposed’ a single technology that all countries in Europe had to collectively adopt. In addition, the ‘one-off’ character of the spectrum allocation process created a ‘now or never’ mindset – licenses had to be obtained in the ‘narrow time window’ that was available or else there would not be a second chance to obtain a 3G license. This created a prisoner’s dilemma for incumbent operators as 3G seemed the only route to safeguard their installed customer base.

Furthermore, firms widely believed that 2G had reached its limits and that, without 3G, the industry would be at a serious disadvantage in serving future customer needs. Firms’ belief in the proposed new technology’s potential was reinforced by various other constituents, such as infomediaries active in sketching “utopian future worlds” (Van Lente, 2000). Various self-reinforcing and mutually dependent forecasts formed part of this rhetoric. It was difficult for consultants, analysts and other experts to make ‘independent’ forecasts and they had to rely on figures from mobile firms. Even inside the mobile companies, the more skeptical views from technologists were swamped by the future visions of brand management and finance (Panzar and Repo, 2005). As the interdependence between analysts’ forecasts and firms’ actions grew, it further reinforced the most optimistic forecasts. Collectively, these forecasts led to a rhetoric space around the proposed new technology as conjectures about future consumer needs turned to imperatives.

The dynamics that emerged are similar to what Van Lente and Rip (1998) have reported about constituents becoming prisoners of an emerging discourse and dominant field vision with little room for voices of dissent. As in the case with mobile communications, while firms craft promising stories about a proposed new technology’s prospects, various institutional actors, such as the media, analysts and even national governments validate these stories in the public domain, conferring legitimacy on them in a self-reinforcing process (Pollock and Rindova, 2003). The collective story of a future generation then gains widespread attention, appearing to be an evident fact and a “scientific canon.” The step from the promise of development of a new technology to the imperative of it happening is readily made and there is really no choice but to push ahead for the new technology (Van Lente, 2000) – a process that contributes to generating “emerging irreversibilities” (Callon, 1991). Being unable to free themselves of the story’s obligations and the collective web spun around the proposed new technology’s potential (Van Lente, 2000; Wagenaar, 1997), the technology’s adoption now seems like the only feasible option as firms decide to make the switch. Based on these observations, we offer the following conjecture for further enquiry.

Conjecture 1. Prospective narratives crafted by focal firms and validated by institutional actors in favor of a proposed new technology generate self-fueling forces for the implementation of the new technology.

5.2. Dampening forces

While some forces can generate momentum for the new technology (Hughes, 1983), others can retard its progress by disrupting the temporal coordination that had been set in place. As we saw in the mobile communications case, lack of specialized 3G handsets and standard battles led to a slowdown in the technology’s adoption. Delays in developing co-specialized assets (Teece, 1986) are likely to arise with any transition as there is not necessarily a one-to-one correspondence between a technology’s future road-map (however conservative it might be) and its implementation over time. Amid future projections and the rush to forge ahead with a new technology, its supporters cannot necessarily factor in all the practical implementation details that will be involved in the development of co-specialized assets for the technology to become operative. Delays in the development of co-specialized assets then lead to delays in the launch and diffusion of the new technology making it possible for the existing technology to catch up. These observations lead to our second conjecture to be explored in other contexts.

Conjecture 2. Delays in the development of co-specialized assets required for the new technology to function retards the diffusion of the new technology and provides an opportunity for the previous technology to catch up.

5.3. Collateral developments

It is understandable that a delay in the creation of co-specialized assets makes it possible for an existing technology to catch up with the proposed new technology. These improvements may come from within the industry but may also result from developments in related industries. For instance, in the context of mobile communications, a “collateral technology” (packet switching) enabled the extension of the old technology (2G), and the extended version (2.5G) adequately addressed the emergent needs of the users around limited use of mobile content services. This delayed the adoption and diffusion of the new technology (3G) even though it offered additional ‘richer’ mobile content services.

The phenomenon of how collateral technologies can revitalize previous generation technologies and mediate transitory paths, albeit to different degrees, has been observed in other domains. For instance, in photo-lithography industry, optical lithography has shown unusual persistence as the dominant manufacturing technology for computer chips since the late 1960s (Henderson, 1995). While the industry was focused on developing “next generation lithography” (NGL) that included technologies such as extreme ultraviolet lithography (EUVL), the discovery of “enhanced liquid immersion” extended the life of optical lithography (Sydow et al., 2007). Infusing optical lithography with a collateral technology required only minor adjustments compared to EUVL for not only extending the path of optical technology but also substantially improving its performance and retarding the adoption of EUVL. Similarly, improvements in steelmaking from the introduction of open-hearth furnaces in the late 1870s enabled the substitution of earlier generation steam boats for sailing ships. Open-hearth furnaces allowed the production of better steel, which in turn enabled boiler plates and boiler tubes to withstand higher pressures. As a result, the earlier generation technology was improved that allowed steam boats to operate more efficiently, and com-

---

12 In enhanced liquid immersion lithography (JUL), a drop of fluid (water or oil) is placed between the optical lens and the wafer. Compared to air, the higher refractive index of fluids leads to a better image resolution. Originally used in microscopy to enlarge the image of the specimen, immersion is used in optical lithography to print miniaturized features onto silicon wafers.
pete more effectively with new generation sailing ships (Dattee, 2007).

In short, collateral innovations in related domains can result in extending the performance of an existing technology thereby closing its performance gap vis-à-vis the new technology. This raises the bar for the new technology that now needs to deliver on additional dimensions to demonstrate its superiority over the old, thereby potentially retarding its adoption and diffusion. Thus, technological developments in related domains also matter in shaping technological transitions. Based on these observations, we offer the following conjecture to be explored in other settings.

**Conjecture 3.** Collateral innovations in related domains can reduce performance gaps between existing and new technological systems, thereby raising the bar for the new technology and retarding its adoption.

5.4. **Emergent shifts in preferences and incentives**

Any closure in the performance gap between the old and the new technology because of forces emerging during the journey again catches the attention of the constituents from the supply, demand and institutional domains. As we noted earlier, during initial stages of new technology emergence, firms craft stories in support of their strategies that institutional actors, such as the media and analysts magnify and legitimize (Pollock and Rindova, 2003; Van Lente, 2000). Now, slippages in the new technology, as we saw in the mobile communications case, may dynamically change expectations of constituencies and their incentives to forge ahead with their plans. For instance, end users may begin to find the old but improved technology acceptable and delay or even refrain from adopting the new technology. In the mobile telephony case, a majority of users continued to use the mobile phone for simple calls and text messages rather than embrace sophisticated services such as making online purchases. Even video telephony – the only application exclusive to 3G – failed to achieve user acceptance.13

These emergent dynamics may problematize the viability of investments during implementation and lead, for instance, to deterring users from different sides of a market. For instance, to generate momentum for the new technology, mobile platforms had to attract ‘complementors’ to develop content for users in the mobile domain. Without compelling content, 3G services were less likely to attract users and achieve critical mass, and without a critical mass, there was little incentive for content providers to develop content for 3G. Such a “chicken and egg” situation, typical of multi-sided platforms (Rochet and Tirole, 2003), may lead to a slowdown in a platform’s growth. Institutional actors, such as infomediaries may then step in and, rather than glorifying the new technology, prevent a smooth transition from one generation to another. Based on these observations, we offer the following conjecture to be explored in other settings.

**Conjecture 4.** As the performance gap between the new and old technologies decreases, the incentives of the various participants to make the transition changes dynamically, thereby shaping the transition process and preventing any one technological system from emerging as a clear winner.

---

13 This had been witnessed much earlier on in the case of fixed line telephones. Video phones have been around since the 1960s but never became popular as people did not want surveillance inside their homes (Lynn et al., 1996).

6. **Conclusion**

We have built on previous work to understand the forces that arise from the interactions between producers, users, and institutional players as they grapple with different elements of technologies during a transition. As we observed in the mobile communications case, new technologies with the promise of higher functionalities do not simply eclipse older ones in a process of creative destruction. Rather, transitions may potentially be ‘creative extensions’: one where elements of existing technologies morph to provide new viable solutions. In the mobile communications case, for instance, the dialectic interplay between forces for change and for continuity led to the emergence of an intermediate technology, 2.5G, that was able to transgress the frame set for 2G and subsume the frame around 3G.

What are the foundations for such a perspective on creative extension? As transitions are necessarily about the future, the future needs to be framed – often dramaturgically with promises and idealizations to raise expectations and attract resources (Lampel, 2001; Van Lente and Rip, 1998). Such framing is accomplished by connecting and coordinating the activities of multiple heterogeneous elements – technologies, suppliers, users and regulators – distributed across time and space (Geels, 2002). This coordination, which has to unfold into an uncertain future, is generated by creating a rhetorical space that engenders the collective mobilization of actors.

However, in making a case for a new technology, the framing of the future requires a “bracketing” of the world, a process that cannot possibly take into account all of the potential “connections, relations and effects” (Callon, 1998:16). Moreover, existing arrangements are likely to be reconfigured during implementation and, consequently, even the most carefully laid out of plans can easily come undone. In addition, framing something new entails contrasting it against what already exists. That is, the present is often problematized in order to make a case for change (Brown et al., 2000). Consequently, as a transition unfolds, those whose interests are aligned with the present may attempt to extend the benefits that constituents may value by continuing with what already exists. In this task, extensions that occur from developments in collateral technologies may come to play a role. And, as in the mobile communications case, user needs, that have yet to emerge and to be tested may congeal around a set of functionalities even though they may fall short of what had been projected earlier. These are all reasons as to why “overflows” (Callon, 1998) will emerge as a transition unfolds, disturbing the anticipated connections between heterogeneous elements that make up a future frame. Because of these overflows, actors’ interests, goals and preferences as well as their interactions with others in a network will be emergent and in a process of continual reconfiguration (Callon, 2008) as transitions unfold through a “series of adaptations” over time (Geels, 2002).

How might various actors in the socio-technical system try to sponsor these transitions? Institutional actors, such as governmental and quasi-governmental actors can play an important role to “steer” transition paths by enacting policies, offering forums for interactions and proactively shaping incentives (Kemp et al., 2001). These dynamics were evident in the role that the Danish government played in coming up with an appropriate mix of policies and incentives to encourage the development and diffusion of wind turbine technology to address Denmark’s energy needs (Garud and Karnøe, 2003). Besides institutional actors, supply-side actors too can play a role in shaping platform dynamics. Specifically, dominant players in privileged positions can engage in “coring strategies” (Gawer and Cusumano, 2008) by managing tradeoffs between controlling key complements while incentivizing third parties to create complementary innovations (Boudreau, 2005, 2006). Others have written about a
“keystone” role (e.g. NTT DoCoMo and its I-mode data-based services for mobile Internet (Peltokorpi et al., 2007)) that firms might play (Iansiti and Levien, 2004) in aligning the interests of multiple constituents so as to generate momentum around a platform.

However, even the best of attempts at managing transitions may come undone because of emergent issues that dynamically change transition parameters. For instance, while institutional actors may attempt to synchronize heterogeneous elements around a new technology, user interests may congeal around a different set of functionalities or value propositions as was the case with Cochlear implants where the profoundly deaf preferred to continue operating within their culture even as the FDA and the NIH attempted to play a steering role (Lane, 1992). Similarly, attempts at gaining platform leadership during a transition may face stiff resistance not only from constituents whose technological turf stands threatened but also by unanticipated developments that may change the rules of the game for the technology sponsor as the transition unfolds. Some of these dynamics were evident in the challenge Sun Microsystems faced from Microsoft’s counter-mobilization strategies as Sun attempted to gain platform leadership by sponsoring its Java software technology (Garud et al., 2002).

Our observations thus place a word of caution against an unrelenting belief by sponsors in the supremacy of a novel technology based on its initial framing and the prevailing wisdom to prematurely abandon old technologies that are sometimes perceived as having reached their normal performance limits (see also Garud and Minuri, 2008). Rather than view transitions as unitary and coherent replacement of existing technologies by new technologies, our findings suggest a conceptualization of transitions as being characterized by the dialectic interplay between the existing and the new as a transition unfolds. In this regard, we see the benefits of fusing insights from recent work on “multi-sided” markets and platforms (e.g. Boudreau, 2006; Evans et al., 2006; Rochet and Tirole, 2003) with notions such as “overflows” (Callon, 1998) and “interpretative flexibility” (the shaping of technologies in use) (Bijker et al., 1987) that focus on emergent developments during a transition. By bringing these literatures together, productive lines of inquiry open around shifting motivations and incentives of various groups involved and how, as a result, an existing technology and its proposed replacement interact during any technological transition.

Acknowledgements

This research has been funded in part by a grant from the Networks Electronics, Commerce and Telecommunications (NET) Institute, www.NETinst.org. We thank the NET Institute for financial support. We also thank the participants for their valuable comments during the ECIS/KSI workshop, May 2008 Eindhoven, The Netherlands, 2008; and the Technology and Innovation Management (TIM) session at the Academy of Management (AOM) conference in Philadelphia, USA, 2007, where an earlier draft of this paper was presented. We thank the anonymous reviewers for Research Policy and Michel Callon for their inputs on an earlier draft of this paper.

References


Pinch, T.J., Bijker, W.E., 1984. The social construction of facts and artefacts: or how the sociology of science and the sociology of technology might benefit each other. Social Studies of Science 14, 399–441.


Seybold, A., 2004. 2.5G or 3G. Get over it. Wireless Week, May 01.


The Economist, 2004a. The mobile phones. Battling for the palm of your hand, April 29.


